

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:41:42 ; Search time 104 Seconds  
(without alignments)  
4387.232 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 4272

Sequence: 1 MWIQVRTMDGROTHVDSLS.....VNQPLQTLVNLQFPFGYNGR 793

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4272	100.0	793	2 Q9P115	Q9P115 homo sapien
2	4263	99.8	793	2 Q96T88	Q96T88 homo sapien
3	3223.5	75.5	829	2 Q7VDF1	Q7VDF1 rattus norv
4	3179.5	74.4	782	2 Q9VDF2	Q9VDF2 mus muscu
5	3175.5	74.3	782	2 Q921H6	Q921H6 mus muscu
6	2929.5	68.6	775	2 Q6PEI0	Q6PEI0 brachydanio
7	2929.5	68.6	775	2 AAH58055	AAH58055 brachydanio
8	2920	68.4	776	2 Q6DRP6	Q6DRP6 brachydanio
9	2384	55.8	803	2 Q7TMI3	Q7TMI3 mus muscu
10	2384	55.8	803	2 AAH60241	AAH60241 mus muscu
11	2379	55.7	803	2 Q8K1T5	Q8K1T5 mus muscu
12	2295	53.7	597	2 Q6IP39	Q6IP39 xenopus lae
13	2295	53.7	597	2 AAH72079	AAH72079 xenopus lae
14	2293.5	53.7	802	2 Q96P04	Q96P04 homo sapien
15	1889.5	44.2	474	2 Q8VTA1	Q8VTA1 mus muscu
16	1442	33.8	516	2 Q8BJP6	Q8BJP6 mus muscu
17	1365.5	32.0	503	2 Q8TAG7	Q8TAG7 mus muscu
18	1207	28.3	299	2 Q8CF71	Q8CF71 homo sapien
19	1005	23.5	189	2 Q9H6S6	Q9H6S6 mus muscu
20	732	17.1	139	2 Q8UJ22	Q8UJ22 homo sapien
21	608	14.2	645	2 Q8VZ02	Q8VZ02 homo sapien
22	600	14.0	641	2 Q9FVS3	Q9FVS3 arabidopsis
23	588.5	13.8	617	2 Q9FKA7	Q9FKA7 arabidopsis
24	588.5	13.8	617	2 Q9FKA7	Q9FKA7 arabidopsis
25	586	13.7	288	2 AAQ65196	AAQ65196 arabidopsis
26	560	13.1	615	2 Q9BZ45	Q9BZ45 homo sapien
27	560	13.1	615	2 Q6NQ90	Q6NQ90 arabidopsis
28	556	13.0	622	2 AAQ65191	AAQ65191 arabidopsis
29	553.5	13.0	765	2 Q9XW58	Q9XW58 oryza sativ
30	533.5	12.5	598	2 Q9C8E0	Q9C8E0 arabidopsis
31	524	12.3	789	2 Q75M36	Q75M36 oryza sativ

32	524	12.3	789	2	AAS88821	AAS88821 oryza sat
33	514.5	12.0	610	2	Q9FW25	Q9FW25 oryza sativ
34	473	11.1	650	2	Q9FVS2	Q9FVS2 arabidopsis
35	427	10.0	461	2	Q81463	Q81463 arabidopsis
36	407	9.5	216	2	Q8BG56	Q8BG56 m mus muscu
37	377	8.8	195	2	Q8BY30	Q8BY30 mus muscu
38	353.5	8.3	432	2	Q22280	Q22280 arabidopsis
39	335	7.8	434	2	Q8H9A3	Q8H9A3 arabidopsis
40	311.5	7.3	299	2	Q8RU61	Q8RU61 deinococcus
41	310.5	7.3	301	2	Q9ADD4	Q9ADD4 streptomyces
42	299.5	7.0	794	1	SUV5_ARATH	SUV5_ARATH arabidopsis
43	279	6.5	684	2	Q7XHM7	Q7XHM7 oryza sativ
44	265.5	6.2	790	1	SUV6_ARATH	SUV6_ARATH arabidopsis
45	264.5	6.2	766	2	Q8H6B0	Q8H6B0 zea mays (m

## ALIGNMENTS

RESULT 1  
Q9P115 PRELIMINARY; PRT; 793 AA.  
ID AC Q9P115;  
DT 01-OCT-2000 (TREMBLrel. 15, Created).  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update).  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update).  
DE Transcription factor ICBP90.  
GN Name=ICBP90;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20110783; PubMed=10646863;  
RA Hopfner R., Mousli M., Jeltsch J.M., Voulgaris A., Lutz Y., Marin C.,  
RA Bellocq J.P., Oudet P., Bronner C.;  
RT "ICBP90, a novel human CCAAT binding protein, involved in the  
RT regulation of topoisomerase I $\alpha$  expression."  
RL Cancer Res. 60:121-128(2000).  
DR EMBL; AF129507; AAF28469.1; -  
DR HSP; Q9UIG0; IF62.  
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . . ; TAS.  
DR GO; GO:0003700; F:transcription factor activity; TAS.  
DR GO; GO:0008283; P:cell proliferation; TAS.  
DR GO; GO:0006357; P:regulation of transcription from Pol II pro. . . ; TAS.  
DR InterPro; IPR011011; FIVE\_PHD\_Znf.  
DR InterPro; IPR003105; G9a.  
DR InterPro; IPR000626; Ubiquitin.  
DR InterPro; IPR001965; Znf\_PHD.  
DR Pfam; PF00628; PHD; 1.  
DR Pfam; PF00240; ubiquitin; 1.  
DR Pfam; PF02182; YDG\_SRA; 1.  
DR PRINTS; PR00348; UBIQUITIN.  
DR SMART; SM00249; PHD; 1.  
DR SMART; SM00184; RING; 2.  
DR SMART; SM00466; SRA; 1.  
DR SMART; SM00213; UBO; 1.  
DR PROSITE; PS00053; UBIQUITIN\_2; 1.  
DR PROSITE; PS01359; ZF\_PHD\_1; UNKNOWN\_1.  
DR PROSITE; PS00016; ZF\_PHD\_2; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; UNKNOWN\_1.  
DR PROSITE; PS00089; ZF\_RING\_2; 2.  
SQ SEQUENCE 793 AA; 89815 MW; D9B4161E892BB014 CRC64;

Query Match 100.0%; Score 4272; DB 2; Length 793;  
Best Local Similarity 100.0%; Pred. No. 2.7e-278;  
Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWIQVRTMDGROTHVDSLSRLTKVEELRRKIQELFHVFGQLQRFYRGKQMGDHTLFD 60  
DB 1 MWIQVRTMDGROTHVDSLSRLTKVEELRRKIQELFHVFGQLQRFYRGKQMGDHTLFD 60

QY	61	YEVRLNDTIQLLVROSILVPHSTKERSLSDTSGCCLGQSEDSKSTHGEAAETDSR	120	Genew; HGNC:12556; UHRF1.
Db	61	YEVRLNDTIQLLVROSILVPHSTKERSLSDTSGCCLGQSEDSKSTHGEAAETDSR	120	GO: GO:0003677; F:DNA binding; IEA.
QY	121	PADEDMWDETELGLYKNEYVDARTNMGAFEAQVVRTRKAPSRDEPCSTSRPALEE	180	GO: GO:0004842; F:ubiquitin-protein ligase activity; IEA.
Db	121	PADEDMWDETELGLYKNEYVDARTNMGAFEAQVVRTRKAPSRDEPCSTSRPALEE	180	GO: GO:0008270; F:zinc ion binding; IEA.
QY	181	DVIYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPDNPKRGFWYD	240	GO: GO:0016567; P:protein ubiquitination; IEA.
Db	181	DVIYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPDNPKRGFWYD	240	GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
QY	241	AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVEFKIERPBGESPMVDPNMRKSGP	300	InterPro; IPR003105; G9a.
Db	241	AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVEFKIERPBGESPMVDPNMRKSGP	300	InterPro; IPR000626; Ubiquitin.
QY	301	SKCHKDDVNRLCRVCACHLGGGRQDPKQLMCDCEDMAFHIYCLDPLSSVSEDEWYC	360	InterPro; IPR001965; Znf PHD.
Db	301	SKCHKDDVNRLCRVCACHLGGGRQDPKQLMCDCEDMAFHIYCLDPLSSVSEDEWYC	360	InterPro; IPR001841; Znf_ring.
QY	361	PECRNDASEVVLAGERLRESKNAKASATSSQSDWGKMACVGRTECTIVPSNHYP	420	Pfam; PF00628; PHD; 1.
Db	361	PECRNDASEVVLAGERLRESKNAKASATSSQSDWGKMACVGRTECTIVPSNHYP	420	Pfam; PF00240; ubiquitin; 1.
QY	421	IPGIPVGTMRFRVQVSGVHRPHVAGIHRGSDGYSILVLAGGYEDDVGHNFFTYTG	480	Pfam; PF02182; YDG_SRA; 1.
Db	421	IPGIPVGTMRFRVQVSGVHRPHVAGIHRGSDGYSILVLAGGYEDDVGHNFFTYTG	480	PRINTS; PR00348; UBIQUITIN.
QY	481	SGGRDLGSKNRTAQSCDQKLTNTNRALNCFAPINDQEGAEAKOWRSKPVVRVNVK	540	SMART; SM00249; PHD; 1.
Db	481	SGGRDLGSKNRTAQSCDQKLTNTNRALNCFAPINDQEGAEAKOWRSKPVVRVNVK	540	SMART; SM00184; RING; 2.
QY	541	GGKNSKYAPAEGRNRYDGIYKVKYWPKEKSGFLVWYLLRRDDDEPGPTWKEGDKRIK	600	SMART; SM00466; SRA; 1.
Db	541	GGKNSKYAPAEGRNRYDGIYKVKYWPKEKSGFLVWYLLRRDDDEPGPTWKEGDKRIK	600	SMART; SM00213; UBQ; 1.
QY	601	LGLTMQYPEGVLEALANRERKENSKEEEOQGGFASPRTGKWKRSAGGPGSRAG	660	PROSITE; PS00553; UBIQUITIN_2; 1.
Db	601	LGLTMQYPEGVLEALANRERKENSKEEEOQGGFASPRTGKWKRSAGGPGSRAG	660	PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
QY	661	SPRRTSKTKTVEPYSLTAQSSLIREDKSNAKLWNEVLASIKDRPASGSPFLSKVEE	720	PROSITE; PS0016; ZF_PHD_2; 1.
Db	661	SPRRTSKTKTVEPYSLTAQSSLIREDKSNAKLWNEVLASIKDRPASGSPFLSKVEE	720	PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
QY	721	TFQCICCOELVFRPITTVCOHNVCKDCLDRSPRAQVFCPCACRYDLGRSYAMQVNPLOT	780	PROSITE; PS00089; ZF_RING_2; 2.
Db	721	TFQCICCOELVFRPITTVCOHNVCKDCLDRSPRAQVFCPCACRYDLGRSYAMQVNPLOT	780	SEQUENCE 793 AA; E65B15657525C89F CRC64;
QY	781	VLNQLFPGYGNGR 793		
Db	781	VLNQLFPGYGNGR 793		
RESULT 2				Query Match 99.8%; Score 4263; DB 2; Length 793;
Q96788				Best Local Similarity 99.7%; Pred. No. 1.1e-27; Indels 0; Gaps 0;
ID	Q96788	PRELIMINARY; PRT; 793 AA.		Matches 791; Conservative 1; Mismatches 1;
AC	Q96788;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Nuclear zinc finger protein Np95.			
GN	Names=UHRF1;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RA	Davenport J.W., Fernandes E.R., Neale G.A.M., Goorha R.M.;			
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF274048; AAK55744.1; -.			
DR	HSSP; Q9UIG0; 1F62.			

```

Db 601 LGLTMQYPEGYLEALANREKENS KREBEQEGGFASPTGKGKWKRSAGGSPSRAG 660
QY 661 SPRTSKTKVPEYSLTAQSSILREKSNKLNWELASIKDRPASGSPQLFLSKVEE 720
Db 661 SPRTSKTKVPEYSLTAQSSILREKSNKLNWELASIKDRPASGSPQLFLSKVEE 720
QY 721 TFCQICCOELVFRPITTVCOHNVCCKDCLDRSFRAQVFCSPACRYDLGRSVMQVQPLQT 780
Db 721 TFCQICCOELVFRPITTVCOHNVCCKDCLDRSFRAQVFCSPACRYDLGRSVMQVQPLQT 780
QY 781 VLINQLFPGYNGNR 793
Db 781 VLINQLFPGYNGNR 793

RESULT 3
Q7TPK1
ID Q7TPK1 PRELIMINARY; PRT; 829 AA.
AC Q7TPK1;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Ac2-121.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Wang G.P., Chai L.Q., Yuan J.Y., Yang K.J.,
RA Yan H.M., Chang C.F., Zhao L.F., Ma H., Wang L., Wang S.F., Han H.P.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY321334; AAP86266.1; -
DR GO; GO:000151; F:ubiquitin ligase complex; IEA.
DR GO; GO:0003677; C:ubiquitin binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003105; G9a.
DR InterPro; IPR000566; Lipocin_cytFABP.
DR InterPro; IPR000626; Ubiquitin.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00628; PHD; 1.
DR Pfam; PF0240; ubiquitin; 1.
DR Pfam; PF02182; YDG_SRA; 1.
DR SMART; SM00249; PHD; 1.
DR SMART; SM00184; RING; 2.
DR SMART; SM00466; SRA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
DR PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR PROSITE; PS50016; ZF_PHD_2; 1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 829 AA; 93222 MW; B5E8327F33FE74BE CRC64;

Query Match 75.5%; Score 3223.5; DB 2; Length 829;
Best Local Similarity 75.1%; Pred. No. 7.8e-208;
Matches 597; Conservative 94; Mismatches 81; Indels 23; Gaps 9;

QY 1 MWIQVMTDGRQHTVDSLSRLTKVLELRKIKIELFHFVEGLQRLFYRGKQMEDGHTLFD 60
Db 56 MWIQVMTDGRKETHVNSLSRLTKVLELRKIKIELFHFVEGLQRLFYRGKQMEDGHTLFD 115
QY 61 YEVRINDTTLQVRLQSLVPLSHPTKERSDSEISDPTDSCCLGQSSDKSSTHGEAAATDSR 120
Db 116 YDVRINDTTLQVRLQSLVPLSHPTKERSDSEISDPTDSCCLGQSSDKSSTHGEAAATDSR 173
QY 121 PADEMDWDETEGLYKVEYVDARDTNMGAMFEAQVVRVTRKAPSRDEPCSSSTRPALEE 180

```

```

Db 174 --DKTWEDTDLGLYKVEYVDVEDNIFGAFRAQVVOVKKALSEEPCCSSIMAPED 231
QY 181 DVIYHVKYDDYPENGVTQMSRDRVARARTIIKWQDLEVGQVVMNLNPNPNKRGFWYD 240
Db 232 DIWHIHKYDDYPENGVDIVAKNVRARARTIIPWEDLEVGQVVMNLNPNPNKRGFWYD 291
QY 241 AEISRKETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGESEPMVDNPMRKSG 299
Db 292 VEICRKQRTARELYGNVLLNDSQLNCRIIIFVDEVFKIELPNEPSPLIGESRRKSG 351
QY 300 PSCKHCKDDVNLRCVCAHLCCGRQDDKQLMCDCECMAFHIYCYLDPPLSSVPSSEDEWY 359
Db 352 PSQCYCKDDENKPCRKACHICGGEAPEKQVLCDCECMAFHIYCYLDPPLTPVPEPEWY 411
QY 360 CPBCRDASEVVLAGELERESKKNKAKASATSSQORDWKGMACVGRTKECTIVPSNHYG 419
Db 412 CPSCRTDSEVVOAGEKLLKSKKAKASATSSRRDWDGKMACVGRKTECTIVPANHFG 471
QY 420 PIPGIPVGTWFRFRVQVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVHGNFTYT 479
Db 472 PIPGVPVGTWFRFRVQVSESGVHRPHVAGIHGRSNDGAYSLVLAGGYEDDVHGNFTYT 531
QY 480 GSGGRDLSGNKRTABOSCDQKLTNNRALLNCFAPINDQEGAEAKDWRSGKPVVRNV 539
Db 532 GSGGRDLSGNKRTABOSCDQKLTNNRALLNCFAPINDQEGAEAKDWRSGKPVVRNV 590
QY 540 KGGKNSKAPABGNRYDGIYKVYKWPBKSGFLWRYLLRRDDDEPGFWTKBGRDKRT 599
Db 591 KGGKNSKAPABGNRYDGIYKVYKWPBKSGFLWRYLLRRDDDEPGFWTKBGRDKRT 650
QY 600 KLGLTMQYPEGYLEALANREKENS KREBEQEGGFASPTGK-GKWKRSAGGSPSR 658
Db 651 QLGLTMQYPEGYLEALANKE--NRKPAKALEG--PSSKIGSKRSKSTGPAFT- 702
QY 659 AGSPRTSKTKVPEYSLTAQSSILREKSNKLNWELASIKDRPASGSPQLFLSKV 718
Db 703 --SP-RVSKSKLEPYTLPLQANLIKEDKGNKLWDVLSLQD-----GPYQIFLSKV 754
QY 719 BETFOCICCOELVFRPITTVCOHNVCCKDCLDRSFRAQVFCSPACRYDLGRSVMQVQPL 778
Db 755 KEAFQCICCOELVFRPITTVCOHNVCCKDCLDRSFRAQVFCSPACRYDLGRSVMQVQPL 814
QY 779 QTVLNQLFPGYNGNR 793
Db 815 QTILNQLFPGYNGNR 829

RESULT 4
Q8VDF2
ID Q8VDF2 PRELIMINARY; PRT; 782 AA.
AC Q8VDF2;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Ubiquitin-like, containing PHD and RING finger domains, 1.
GN Name-Uhrf1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,

```

Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.M., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Faney J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
Krzyszynski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
Jones S.J., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";

RT  
PROC. Natl. Acad. Sci. U.S.A.  
[2]  
RN  
SEQUENCE FROM N.A.  
RP  
STRAIN=CZECH II;  
RC  
TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;  
RA  
Strausberg R.;  
to the ENZYME/Genbank/EMBL databases.

RL Submitted (JAN-2002) to the EMBL/Genbank/ DDBJ  
DR EMBL; BC022167; AAH22167.1; -.  
na HSSP: O9UIG0: 1F62.

DR MGD; MGI:1338889; Uhrf1.  
DR GO; GO:0000151; C-ubiquitin ligase complex; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0004682; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO:0008270; P:zinc ion binding; IEA.  
DR GO:0016567; P:protein ubiquitination; IEA.  
DR GO:0001655; P:regulation of transcription, DNA-dependent; IEA.  
DR GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR011011; FYVE\_PHD\_Znf.  
DR InterPro; IPR003105; G9a.  
DR InterPro; IPR003105; G9a.

DR InterPro; IPR000566; Lipocalin cyclase.  
DR InterPro; IPR000626; Ubiquitin.  
DR InterPro; IPR001965; Znf PHD.

DR InterPro; IPR001841; Znf\_ring.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00628; PHD; 1.  
DR Pfam; PF00628; PHD; 1.

DR Pfam; PF00240; ubiquitin; 1.  
DR Pfam; PF02182; YDG SRA; 1.  
DR Pfam; PF00240; ubiquitin; 1.

DR PRINTS; PR00348; PHD; 1.  
DR PRINTS; PR00348; PHD; 1.  
DR SMART; SM00249; PHD; 1.  
DR SMART; SM00184; RING; 2.

DR	SMART	SM00466	SRA	1.
DR	SMART	SM00466	SRA	1.
DR	SMART	SM00213	UBQ	1.

DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_1.  
DR PROSITE; PS00253; UBIQUITIN\_2; 1.  
DR PROSITE; PS00503; UBIQUITIN\_2; 1.  
DR PROSITE; PS01350; 7S BDN 1; UNKNOWN\_1.  
DR PROSITE; PS01350; 7S BDN 1; UNKNOWN\_1.

DR PROSITE; PS00559; ZF\_RING\_2; UNKNOWN 1.  
DR PROSITE; PS00560; ZF\_PHD\_2; 1.  
DR PROSITE; PS00561; ZF\_RING\_1; UNKNOWN 1.

```
DR PROSITE; ZF_RING_2; 2.  
SQ SEQUENCE 782 AA; 88319 MW; DC5EEDFCDF779074 CRC64;
```

Query Match 74.4%; Score 3179.5; DB 2; Length 782;  
Best local similarity 73.6%;  
Pred. NO. 6.6e-205;

Matches 590; Conservative 89; Mismatches 94; Indels 29; Gaps

1 MWIQVRIMDGRQTHTVDSLSRLTKRVEELRRKIQELHFHVEFGDQKLF INGAQVLECHILIT

DB I MWIQVRINDGNEIHI VASLSKAIK QEEAAETDTSR  
61 YEVRNPTIQLLVRSIVLPSTKRDSELSDTDSGCCLGQSESDKSSSTHGEAAETDSR

61 YDVRNDTIQLLVQRSLALPLSTKERDSELSDSGCVGVGHSESKSSTHGEAAEAD--

QY 121 PADEDMDTELGLYKNEYVDARDTNMGAWFEAQVVRVTRKAPSRDEPCSSTSRPALEE  
|. . . . .| : : : : :| : : : : :| : : : : :|

Db 119 --DKTWEDTDGLGYKVNEYVDVRDNI FGAWFAQVVQKRALSEDEPCSSSAVKTSSE

[illegible]

DB 177 DIMIHVKIDDIPFERGVDI VAGIUVAGGKAVLZKNERFQ  
241 ARISKRRETRARELYANV-VLGDDSLNDCRIIFVDEVFKIERPGEPSMVDNPMR

237 VEICRKQRTARELYGNIRLNDSQLNNCRIMFVDEVLMIELPKERRELIASPSQPPFA

296	QY	----	RKSGPSCKHCKDDVNRLCRVCA	CHLCCGGRQDPDKQLMCECD	ECDMAPHIYICLDDPLPS	351
297	Db	----	LNTGKSGSCRFCKDDENK	PKCKACHVCGGREAPEKQL	LDCEDMAFHLYCLKPPLTS	356
352	QY	----	VPSEDEWYCEPCRNDA	SEVVVLAGERLRESKKNAKMASAT	SSSSORDWKGMACVGRTEKCT	411
357	Db	----	VZPEPEWYCPSCRT	SDSSSEVVVAGEKLKESKKAKMASAT	SSSSRRDDWKGMACVGRTECT	416
412	QY	----	IVPSNHYGPIGTPIVCT	WMRFRVOVSESGVHRPHVAGIHGR	SNDGSGSYSLVLAGYEDD	471
417	Db	----	IVPANHFGBIPGVGV	CTMMRFRVOVSESGVHRPHVAGIHGR	SNDGAGSYSLVLAGYEDD	476
472	QY	----	HGNFFTYTSGGGRD	LSGNKRTAEQSCDQKLTNNRAL	NCFAPINDOEAGAKDWRSKG	531
477	Db	----	NGNYFTYTGSGGRD	LSGNKRTAGQSDQKLTNNRAL	NCCHSPIN-EKGAEDWQKG	535
532	QY	----	PVRVWRNVKGKNSK	YAPAGSNRYDGLYKVVKYVPEK	GKSGFLVWYLLRRDDDDPGPWT	591
536	Db	----	PVRVWRNMKGKHS	KYAPAGSNRYDGLYKVVKYVPER	GKSGFLVWYLLRRDDTEPEWT	595
592	QY	----	KEGKDRITKILGLT	WQYPEGYLEALANERKENS	KREBEEOEGGFASPRTKGKWKRS	651
596	Db	----	REGKDRITRQLGLT	WQYPEGYLEALAN-----	KEKSRKRPAKALEOGPSSSKTGKQK	648
652	QY	----	AGGPSPSAGSPR	RTSKTKTKVEPYSYLAQOSSL	IREDKSNAKLWNEVLASLKD	RPSAGSPFF 711
649	Db	----	-STGPT-LSSP-R	ASKSKSLPEYTLSEGOANL	IHKDGNAKLWDDVLTSLQD	-----GPY 700
712	QY	----	QLFLSKVEETFOCH	CCOELVEPRTTTCQHNWCKD	CLDRSPRAQVFCPACHYDLGRSYA	771
701	Db	----	QIFLSKVKEAFQC	ICCCOELVEPRTTTCQHNWCKD	CLDRSPRAQVFCPACHYDLHSSP	760
772	QY	----	MQVNPQLOTVLN	QNLFPYGNGR 793	----	----
761	Db	----	TRVNQPILOTILN	QNLFPYGVNGR 782	----	----

## RESULT 5

ID	Q9Z1H6		PRELIMINARY;	PRT;	782 AA.
AC	Q9Z1H6;				
DT	01-MAY-1999	(TrEMBLrel. 10,	Created)		
DT	01-MAY-1999	(TrEMBLrel. 10,	Last sequence update)		
DT	05-JUL-2004	(TrEMBLrel. 27,	Last annotation update)		
DE	Nuclear protein np95	(Nuclear zinc finger protein Np95).			
OS	Name=Uhrfl; Synonyms=Np95;				
GN	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata;	Cranialta; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia;	Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxId=10090;				
[1]					
RN	SEQUENCE FROM N.A.				
RP	TISSUE=Pre-Tcell;				
RC	MEDLINE=99099250; PubMed=9880673;				
RA	Fujimori A., Matsuda Y., Takemoto Y., Hashimoto Y., Kubo E., Araki R.,				
RA	Fukumura R., Mita K., Tatsumi K., Muto M.;				
RT	"Cloning and mapping of Np95 gene which encodes a novel nuclear				
RL	protein associated with cell proliferation.";				
RL	Mamm. Genome 9:1032-1035(1998).				
[2]					
RN	SEQUENCE FROM N.A.				
RP	Davenport J.W., Fernandes E.R., Neale G.A.M., Goorha R.M.;				
RL	Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.				
DR	ENBL; D87908; BAA74579.1; -.				
DR	ENBL; AF274046; AAK55743.1; -.				
DR	HSPSP; OSUIG0_1F62.				
DR	MGD; MGI:133889; Uhrfl.				
DR	GO; GO:0005634; C:nucleus; IEA.				
DR	GO; GO:0000151; C:ubiquitin ligase complex; IEA.				
DR	GO; GO:0003677; F:DNA binding; IEA.				
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.				
DR	GO; GO:0008270; F:zinc ion binding; IEA.				





## RESULT 7

```
QY 414 PSNHYGPVGPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSYSLVLAGGYEDVDHG 473
Db 411 PSNHYGPVGPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSYSLVLAGGYEDVDHG 470
QY 474 NFFTYTSGGGRDLGSKNRTABQSCDQKLTNNRALLALNCFAPIINDQBAKADWRSKPV 533
Db 471 NEFTYTGSGGRDLGSKNRTABQSCDQKLTNNRALLALNCFAPIINDQBAKADWRSKPV 530
QY 534 RVVRNVKGGKNSKYAPAEGRNRYDGIYKVKYKWEKSGFLVWRYLLRRDDBPFWTKE 593
Db 531 RVVRNSKGRKHSKYSPEDGNRYDGIYKVKYKWEKSGFLVWRYLLRRDDBPFWTKE 590
QY 594 GKDRIKKGLTMOYPEGYLEALANRERKENSREBEEOQEGGFASPRTKGKWKKSAG 653
Db 591 GKDRIKKGLTMOYPEGYLEALANRERKENSREBEEOQEGGFASPRTKGKWKKSAG 650
QY 654 GPPSRAGSPRR-TSKTKTVEPYSLTAAQSSLIREDKSNALNVEVLASLKDREPASGPFQ 712
Db 642 -SMEEKSSPTKGTPTKMKVAYKLSKEQKALIKDELNKKLWDEAMESLSLGR- 694
QY 713 LFLSKVEETFOCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCACRYDLGRSYAM 772
Db 695 -FVNKVEEVFLCICCOEVVYQPIITTECOHNVCKDCLDRSFRAQVFCPCACRYDLGRSYAM 753
QY 773 QVNPQLQVTLNOLFPYGYNGR 793
Db 754 AVNKPQLQVTLNOLFPYGYNGR 774

RESULT 8
Q6DRP6
ID Q6DRP6 PRELIMINARY; PRT; 776 AA.
AC Q6DRP6;
DT 01-OCT-2004 (TremBLrel. 28, Created)
DT 01-OCT-2004 (TremBLrel. 28, Last sequence update)
DE NP95.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Amsterdam A., Nissen R.M., Sun Z., Swindell E.C., Farrington S.,
RT Hopkins N.;
RL "315 Genes Essential for Early Zebrafish Development.";
DR Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2004).
SQ SEQUENCE 776 AA; 87170 MW; 87FE9FC7A1F8664C CRC64;

Query Match 68.4%; Score 2920; DB 2; Length 776;
Best Local Similarity 66.8%; Pred. No. 1.8e-187;
Matches 536; Conservative 113; Mismatches 117; Indels 36; Gaps 12;

QY 1 MWIQRVMDGRTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQWEDHTLFD 60
Db 1 MWIQRVMDGRTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQWEDHTLFD 60
QY 61 YEVRINDTIQLLVRSQSLVPLSTKVERSELSDTDSGCCLGQSESDSKSTHGEAAAEATDSR 120
Db 61 YNVGLNDIVQLLVQAATAVLPKDKAEALSDDSGGSAQSESDSKSTHGEAAAEATDSR 120
QY 121 PADEMDMD -ETELGLYKVEYVDARTNMCAWFAQVVRVTRKAPSDPECSSTSRPAL 178
Db 121 SGQTDATDLIDPGFGFYKINEFVDARDLNMCAWFAQVVRVTRKAPSDPECSSTSRPAL 171
QY 179 BEDVIYHKVDYDPENGVMQNSRDVRAARTIIKQDLEVGQVVMVLYNPDNPKERGF 238
Db 172 -ESIVYHKVEDYDPENGVMQNSRDVRAARTIIKQDLEVGQVVMVLYNPDNPKERGF 230
QY 239 YDAHSKRRTRETRARELYANVVLGD -DSLNDRCRIIFVDFVKTERPG -EG-SPWVDNP 293
Db 239 YDAHSKRRTRETRARELYANVVLGD -DSLNDRCRIIFVDFVKTERPG -EG-SPWVDNP 293
```

```
Db 231 YDAHSKRRTRETRARELYANVVLGD -DSLNDRCRIIFVDFVKTERPG -EG-SPWVDNP 290
QY 294 MRKKS-GPSCKHCKDQVNLRCACACACACACACACACACACACACACACACACACACAC 352
Db 291 LKQSNGPECKVCKDQVNLRCACACACACACACACACACACACACACACACACACACAC 350
QY 353 PSDEWYCPEDRNDASEVVLAGEKESKKAASASSSSQSDQSDQSDQSDQSDQSDQSDQSD 412
Db 351 PDDEDWYCPEDRNDASEVVLAGEKESKKAASASSSSQSDQSDQSDQSDQSDQSDQSDQSD 410
QY 413 VPSNHYGPVGPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSYSLVLAGGYEDVDHG 472
Db 411 VPSNHYGPVGPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSYSLVLAGGYEDVDHG 470
QY 473 GNFFTYTSGGGRDLGSKNRTABQSCDQKLTNNRALLALNCFAPIINDQBAKADWRSKPV 532
Db 471 GNFFTYTSGGGRDLGSKNRTABQSCDQKLTNNRALLALNCFAPIINDQBAKADWRSKPV 530
QY 533 RVVRNVKGGKNSKYAPAEGRNRYDGIYKVKYKWEKSGFLVWRYLLRRDDBPFWTKE 592
Db 531 RVVRNSKGRKHSKYSPEDGNRYDGIYKVKYKWEKSGFLVWRYLLRRDDBPFWTKE 590
QY 593 EGKDRIKKGLTMOYPEGYLEALANRERKENSREBEEOQEGGFASPRTKGKWKKSAG 652
Db 591 EGKDRIKKGLTMOYPEGYLEALANRERKENSREBEEOQEGGFASPRTKGKWKKSAG 650
QY 653 GPPSRAGSPRR-TSKTKTVEPYSLTAAQSSLIREDKSNALNVEVLASLKDREPASGPFQ 711
Db 643 -SMEEKSSPTKGTPTKMKVAYKLSKEQKALIKDELNKKLWDEAMESLSLGR- 695
QY 712 LFLSKVEETFOCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCACRYDLGRSYAM 772
Db 696 -FVNKVEEVFLCICCOEVVYQPIITTECOHNVCKDCLDRSFRAQVFCPCACRYDLGRSYAM 753
QY 773 QVNPQLQVTLNOLFPYGYNGR 793
Db 754 AVNKPQLQVTLNOLFPYGYNGR 775

RESULT 9
Q7TMI3
ID Q7TMI3 PRELIMINARY; PRT; 803 AA.
AC Q7TMI3;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DE NP95-like ring finger protein.
GN Name=NIRF; Synonyms=Uhrf2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Swiss Webster/NTH; TISSUE=Normal whole embryo;
RX PubMed=14741369;
RA Mori T., Li Y., Hata H., Kochi H.;
RT "NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a
RL PEST-containing nuclear protein.";
FEBS Lett. 557:209-214(2004).
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Diatchenko L., Marusina K., Moore T., Max S.I., Wang J., Hsieh F.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Iqbal N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```



```
Query Match          55.8%; Score 2384; DB 2; Length 803;
Best Local Similarity 54.3%; Pred. No. 1.9e-151;
Matches 456; Conservative 122; Mismatches 178; Indels 84; Gaps 13;

OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukemic cell line;
RA Davenport J.W., Fernandes E.R., Neale G.A.M., Goorha R.M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AF274047; AAM33798.1; -.
DR HSP; Q9UIG0; I1F62.
DR MGD; MGI:1923718; Uhrf2
DR GO; GO:0000451; C:ubiquitin ligase complex; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011011; FYVE_PHD_Znf.
DR InterPro; IPR003105; G9a.
DR InterPro; IPR006626; Ubiquitin.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00628; PHD; 1.
DR Pfam; PF02182; YDG_SRA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00249; PHD; 1.
DR SMART; SM00184; RING; 2.
DR SMART; SM00466; SRA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS0053; UBIQUITIN_2; 1.
DR PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR PROSITE; PS0016; ZF_PHD_2; 1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 803 AA; 90091 MW; 99C82799305F3FA0 CRC64;

Query Match          55.7%; Score 2379; DB 2; Length 803;
Best Local Similarity 54.2%; Pred. No. 4.1e-151;
Matches 455; Conservative 122; Mismatches 179; Indels 84; Gaps 13;

QY 1 MWIQRVMDGRQTHVDSLSRLTKVEELRRKIQELFVPEGLQRLFYRGKQMEDGHTLFD 60
DB 1 MWIQRVTDGSGTQRTIEDVSRKATIEELRERWALFDVREPCQLFYRGKQLENGYTLFD 60
QY 61 YEVRNDTIQLVRSQSLVPHSTKERDSELSLTDGCGCLGQSESDKSTHG-----EAA 114
DB 61 YDVGLNDIIQLLRPDSLSTKQNDQV-----KPSHNPVKVKTKAR 105
QY 115 AETDSRPAD--EDMWDETELGLYKVNVEYDARDTNMGAFEAQVVRVTR-----KAP 164
DB 106 GSSSSQPSARTCLIDPGLYKVNELVDARDVGLGAFEAHISVTRASGHRGKTP 165
QY 165 SRD-----EPCSTSRPALEEDVIYHYKDYDPENGVMQ 199
DB 166 LKNGSSYKRTNGNVNHNKENTKLDNVPSTNSDSVADEVDIYHIEYDEYESSILEM 225
QY 200 NSRDVRAARTILKQDLEVGQVVMNPNPNKRGFWYDAEISR-KRETRTARELYAN 258
DB 226 NVKDLRPRARTILKWNELNVGVVMNPNVENPKRGFWYDAEITTLKTSRTKKEVRVK 285
QY 259 VVLG--DDSLNDCRIIFVDEVFIERPGEKSPM--VDNPMRRKSGPSCKHCKDDVNLRCR 314
DB 286 VFLGSGEGTLCNCRVMSVDEIFKIEKPG-AHPISFADGKFLRKNDPECDLGGDPDKTCH 344
QY 315 VCACHLCGRQDDPKQLMCDCEMAFIYCLDPLSPVSEDEWYCEPCRNDASEVVLG 374
DB 345 MCSCHKCKEGRDNNQLLDCENMAYHIYCLSPDLKVPDEEYWCPSCKTDSSEVVKAG 404
QY 375 ERLRESKKNAKASATSSQSDQDWGKMACVGRTECTIVPSNHYGPIPGVGTWVRFRV 434
DB 405 ERLKSKKAKMPASASTESRRDWRGMACVGRTECTIVPSNHYGPIPGVGTWVRFRV 464
QY 435 QVSEGVHPRPHVAGIHGRNDGYSYLVLAGGEDVDYDGHGNTFFYTGSGGDLISGNKRTAE 494
DB 465 QVSEAGVHPRPHVGGIINGRNDGYSYLVLAGGEDVDYDGHGNTFFYTGSGGKNLGNKRIGA 524
QY 495 QSCDQKLTNTRALNCFAPINDQGEAKDWRSGKPVVRVNRVNGKNSKYAPAEGRN 554
DB 525 PSADQTLTNMRALNCDAPLDDKIGAERNWRACKPVVRVIRSPKRLSKYAPAEGRN 584
QY 555 YDGIYKVKYWPKEGKS-GFLVWRYLLRRDDDEPGWTKGDKRIKKLGLTWQYPEGYLE 613
DB 585 YDGIYKVKYWPKEISSHGFLVWRYLLRRDDVEPAPWTSEGIERSRLCLRQYPAGY-- 642
QY 614 ALANRERERKENSKEEEEOEGGFASPRTGKWKRKSGAGGSPRSGSPRRTSKTKVPE 673
DB 643 -----PSEKEGKTKGQSKQSEATKPPA-----SDCEPGDSKVLKASDSTDAVEA 690
QY 674 YSLTAQSSLIREDKSNAKLNVEVLASLKDREPASGPPFOLFLSKVEETFOCICCOELVPR 733
DB 691 FOLTPOQRLIREDQONQKLWDELVASLVEGEN-----FLKLEQSFMCVCCOELVYQ 743
QY 734 PLTTVCQHNVCCKDLDRFRAQVFCSPACRYDLGRSYAMQVNPLOTVLNQLFPQYGNR 793
DB 744 PVTTECFHNVCKDCLQRSFKAQVFCAPCRHDLGQNYVMVLNETLQTLIDLFPFGYSNGR 803

RESULT 11
Q8KLI5
ID Q8KLI5 PRELIMINARY; PRT; 803 AA.
AC Q8KLI5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nuclear zinc finger protein Np97.
GN Name=Uhrf2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```





```

ID AAH72079 PRELIMINARY; PRT; 597 AA.
AC AAH72079;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.U., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Glimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smalish D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC072079; AAH72079.1; -.
KW Hypothetical protein.
FT NON_TER 597
SQ SEQUENCE 597 AA; 67054 MW; 273BEC791D9FA86E CRC64;

Query Match 53.7%; Score 2295; DB 2; Length 597;
Best Local Similarity 70.2%; Pred. No. 1.2e-145;
Matches 425; Conservative 74; Mismatches 92; Indels 14; Gaps 8;

QY 1 MWIQRVMDGROTHVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMGDHTLFD 60
DB 1 MWIQRVMDGROTHVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMGDHTLFD 60
QY 61 YEVRLNDTQLAVRSLVLPHS--TKERDSELDSDTSGCLQSGSDSKSTHGEAAETD 118
DB 61 YSVGLNDIVQLVRQ---IPDSFPPTKHCELSASAGCGSGQRDSGSG--GEGAMDVD 115
QY 119 SRPAEDMMDETELGKLVKNEVVDARDTNGAWFAEQVVRTRK-APSRDEP-CSSTSRP 176
DB 116 GQ-SIIGENVGTSLYKQNDLVARDLNMGAWFAEQIVNKKVGPVGTLPDEVSDTS-- 172
QY 177 ALEEDVIVHKYDDYPENGVMQNSRDVRARARTIKQDLQEVGVQVMVNPDPNPKERG 236
DB 173 VTSDAIIYHKVEDYPENGVVQLTCKDVLRLARTLTPWHEIKVGQVMVNPDPNPKERG 232

```

```

QY 237 FWDAEISRXRETRTARELYANVLGD--DSLNDCRIFVDEVEFKIERPGESEPMVDNPM 294
DB 233 YWYDAEILRKHEKSKIKIYAKVLLGDAGDSLNDCRIFVNEIYKIEEPGSTYLTNTSPQ 292
QY 295 RRSKSGSCHEKCKDDVNLCHVCHLCGGRDDPKQKMLCMDECDMAFIYCLDPLSSVPS 354
DB 293 KRONGPECKHCKDNPKRACSMCACCTCGGKQDPEKQLLCECDLAFHIYCLKPLSLVIPQ 352
QY 355 EDWYCYPECNDASEVVLAGERLESKNAKMASATSSSSORDWKGKMACVGRTEKCTIVP 414
DB 353 DEDWYCPDCNDASEVVLAGEKLESKNAKMASANSSORDWKGKMACVGRRECTIVP 412
QY 415 SNHYGPIPGIPVGTWVRFRVQVSESGVHRPHVAGIHGRSNDGYSYSLVLAGGYEDDVHGN 474
DB 413 SNHYGPIPGVPGVGTWVKFRVQVSESGVHRPHVAGIHGRSNDGYSYSLVLAGGYEDDVGN 472
QY 475 FFFYTGSGGRDLSCNKRRTAEQSCDQKLTNTRALALNCFAPINDQGAEAKQWRSGKPV 534
DB 473 EFTYTSGGRDLSCNKRRTAEQSCDQKLSNMNRALALNCFAPINDQGAEAKQWRSGKPV 532
QY 535 VVRNVKGGKNSKYAPAEAGNRYDGIYKVYKWPBKSGFLVWRYLLRRDDDEPGWTKEG 594
DB 533 VVRNVKGGKNSKYAPAEAGNRYDGIYKVYKWPBKSGFLVWRYLLRRDDDEPGWTKEG 592
QY 595 KRIK 599
DB 593 KKKKK 597

RESULT 14
Q96PU4 PRELIMINARY; PRT; 802 AA.
AC Q96PU4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NP95-like ring finger protein (Nuclear zinc finger protein Np97).
GN Name=NIRF; Synonyms=UHRF2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22165473; PubMed=12176013;
RA Mori T., Li Y., Hata H., Ono K., Kochi H.;
RT "NIRF, a novel RING finger protein, is involved in cell-cycle
RT regulation.";
RT Biochem. Biophys. Res. Commun. 296:530-536 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=14741369;
RA Mori T., Li Y., Hata H., Kochi H.;
RT "NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a
RT PEST-containing nuclear protein.";
RL FEBS Lett. 557:209-214 (2004).
RN [3]
RP SEQUENCE FROM N.A.
RA Davenport J.W., Fernandes E.R., Neale G.A.M., Goorha R.M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AB071698; BAB68317.1; -.
DR EMBL; AF274049; AAM33799.1; -.
DR HSP; Q9UIG0; 1F62.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011011; FYVE_PHD_Znf.
DR InterPro; IPR003105; G9a.

```

DR	InterPro; IPR000626; Ubiquitin.	
DR	InterPro; IPR001965; Znf_PHD.	
DR	InterPro; IPR001841; Znf_ring.	
DR	Pfam; PF00628; PHD; 1.	
DR	Pfam; PF00240; ubiquitin; 1.	
DR	Pfam; PF02182; YDG_SRA; 1.	
DR	Pfam; PF00097; zf-C3HC4; 1.	
DR	SMART; SM00249; PHD; 1.	
DR	SMART; SM00184; RING; 2.	
DR	SMART; SM00466; SRA; 1.	
DR	SMART; SM00213; UBO; 1.	
DR	PROSITE; PS00053; UBIQUITIN_2; 1.	
DR	PROSITE; PS01359; ZF_PHD_2; UNKNOWN_1.	
DR	PROSITE; PS00016; ZF_PHD_2; 1.	
DR	PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.	
DR	PROSITE; PS00089; ZF_RING_2; 1.	
KW	Metal-binding; Zinc; Zinc-finger.	
SQ	SEQUENCE 802 AA; 89984 MW; 190E26D5A347A7FA CRC64;	
Query Match 53.7%; Score 2293.5; DB 2; Length 802;		
Best Local Similarity 53.3%; Pred. No. 2.3e-145;		
Matches 444; Conservative 126; Mismatches 192; Indels 71; Gaps 14;		
QY	1 MWIQRVTMDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60	
DB	1 MWIQRVTMDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60	
QY	61 YEVRNDTIQLVRSQSL-VLPHSTKERSLSLSDTSGCCLGQSSDKSSTHGEAAAEYDS 119	
DB	61 YDVGNDIQLVRSQSL-VLPHSTKERSLSLSDTSGCCLGQSSDKSSTHGEAAAEYDS 114	
QY	120 RPAEDMDETELGLYKNEVDARDTNMGAWFEAQVVRVTR-----KABSD----- 167	
DB	115 ARA---RLIDPFGFYGIYKNEIVDARDVGLGAWFEAHVSHVTRASDQSGRGKTPLNKGSSC 171	
QY	168 -----EPCSSSRP---ALEEDVIYHVYDDYPENGWVQMRSDVRA 206	
DB	172 KRTNGNIKHKSKENTKLDSPVTSNSDCVADEVIYHIQDEYFESGTLNMVNDLRP 231	
QY	207 RARTIKQDLEVGQVMVLYNPNPKRGFWYDAISR-KRETTARELYANVVLG--D 263	
DB	232 RARTILKWNELNVGDVVMVNVESPGQGFWDARITTLKTSRKKELRVKIFLGSE 291	
QY	264 DSLNDCRIIFVDEFEKIPRREGSPM--VDNPMRKSGPSCKHCKDDVNRCLRCVCAHLG 321	
DB	292 GTLNDCKIISVDEIFKIRBP--AHLPLSFADGKFLRRNDPECDLGGDPEKKCHSCSRVC 350	
QY	322 GGRDDPKQLMCDCEMAFYICLDPPLSPSEDEWYCPCRNDASEVYLALERLESK 381	
DB	351 GKGHEPNQLLCECNVAHYIYCLNPPLDKVPEEYVYCPSCKTDSSSEVVKAGERLKXSK 410	
QY	382 KNAKMASATSSQRDWMGKMACVGTKECTIVPSNHYGPIPGIPVGTWRRFRVQVSEGV 441	
DB	411 KKAKMPSASTESRRDWMGRMACVGTRECTIVPSNHYGPIPGIPVGTWRRFRVQVSEAGV 470	
QY	442 RHPVAGTHGSDGYSILVLAGGYEDVDHGNFFTYTSGGRDLSGNKRTAEQSCDOKL 501	
DB	471 RHPVGGHGRSNGAYSLVLAGGFADEVDRDDEFTYTGSGGNLAGNKRIGAPADQTL 530	
QY	502 TMTNEALALNCFAPINDOGEABAKDWRSGKPVVRVNRVNGGKNSKYAPAEGRNYDIYKV 561	
DB	531 TMTNEALALNCDAPLDDKIGAESRNRWAGKPVVRVIRSPFAGRKISKYAPAEGRNYDIYKV 590	
QY	562 KYVWEKGS--CFLVWRVYLLRDDDEPGTWTEGDKRIKKLGLTWQYEGYLEALANRER 620	
DB	591 KYVWEIISSSHGFLVWRVYLLRDDDEVPAPWTSEGIERSRLRLRLQYPAGY-----PSDK 645	
QY	621 EKENSREBEEQOEGFASPRGTGKWKRSAGGSPSRAGSPRRTSKTKVPEFYSILTAQ 680	
DB	646 EGKPKGQSKKQPSGTTKRP-----ISDDCPSASKVYKASDAEIAEAFQITPQ 696	
QY	681 SSLIREDKSNALWNEVLASLKDORPASGSPFOLFSLKVEETFCQICCBELVFRPRTTVCQ 740	

Query Match 44.2%; Score 1889.5; DB 2; Length 474;  
Best Local Similarity 72.8%; Pred. No. 1.6e-118;  
Matches 348; Conservative 58; Mismatches 59; Indels 13; Gaps 3;

QY	1 MWIQRVTMDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60	
DB	1 MWIQRVTMDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60	
QY	61 YEVRNDTIQLVRSQSL-VLPHSTKERSLSLSDTSGCCLGQSSDKSSTHGEAAAEYDSR 120	
DB	61 YDVRNDTIQLVRSQSL-VLPHSTKERSLSLSDTSGCCLGQSSDKSSTHGEAAAEAD-- 118	
QY	121 PAEDMDWDETELGLYKNEIVDARDTNMGAWFEAQVVRVTRKAPRDEPCSSTSRPALEE 180	
DB	119 --DKTWEDTDLGLYKNEIVDVRDNIFGAWFEAQVVRVQVQKRALSEDEPCSSAVKTS 176	
QY	181 DVIYHVYKDDYDYPENGWVQMRSDVRAARTIKKQDLEVGQVMVLYNPNPKRGFWYD 240	

```

Db 177 DIMVHVXYDDYPHGVDDIVKAKNVRARARTVWPENLKVGQVVMANYNDYPRKEGFWD 236
QY 241 AEISRKRETRARELYANV-VLGDDSLNDCRIIFVDEVFKTERPEGESEPMVDNPMR---- 295
Db 237 VEICRKEQTRARELYGNIRLNLNDSQLMNCRIMFVDEVLMIELPKERRPLIASPSQPPPA 296
QY 296 ----RKSGPCKHCKDDVNRLCRYCACHLGGRODDKQLMCDCECDMAFHLYCLDPPLSS 351
Db 297 LRNTGKSGPSCRFCKDDENKFCRKCACHVCGGREAPEKQLLCECDMAFHLYCLKPPLTS 356
QY 352 VPSEDEWYCPCECRNDASEVVLAGERLRESKKNAKMASATSSQRDWMGKMACVGRTECT 411
Db 357 VPPEPEWYCPSCRTDSSEVVQAGEKLESKKAKMASATSSSRDWMGKMACVGRTECT 416
QY 412 IVPSNHYGPIPGIPVGTWRRFRVOVSESGVHRPHVAGIHGRSNDGSYSVLVLAGGYEDD 469
Db 417 IVPANHFGRIPGVFVGTWRRFRVOVSESGVHRHLVAGIPWRSNDGAYSLVLAGGYEDD 474

```

Search completed: November 1, 2004, 15:45:15  
Job time : 108 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:41:42 ; Search time 28 Seconds  
(without alignments)  
2724.996 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 4272

Sequence: 1 MWIQVTRMDGRQTHTVDSLS.....VNQPLQTVLNQLPFGVGNR 793

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: Pirt1:\*

2: Pirt2:\*

3: Pirt3:\*

4: Pirt4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	600	14.0	641	2 E96612	probable transcrip
2	556	13.0	622	2 H96684	probable RING zinc
3	533.5	12.5	598	2 A96685	probable RING zinc
4	473	11.1	650	2 D96612	hypothetical prote
5	427	10.0	461	2 T01825	hypothetical prote
6	353.5	8.3	432	2 T00949	hypothetical prote
7	311.5	7.3	299	2 C75384	conserved hypothet
8	299.5	7.0	794	2 D84765	similar to mammali
9	265.5	6.2	788	2 C84616	similar to mammali
10	229.5	5.4	650	2 T06648	hypothetical prote
11	226	5.3	651	2 F84743	hypothetical prote
12	194.5	4.6	669	2 F96756	similar to mammali
13	192.5	4.5	954	2 G86312	hypothetical prote
14	177.5	4.2	4957	2 T03455	hypothetical prote
15	177.5	4.2	5262	2 T03454	hypothetical prote
16	176	4.1	1829	2 T34239	ALR protein - huma
17	174.5	4.1	811	2 T08738	hypothetical prote
18	162.5	3.8	1280	2 T51500	hypothetical prote
19	158	3.7	371	2 A55302	hypothetical prote
20	157	3.7	1787	2 T20160	probable transcrip
21	155.5	3.6	1518	2 D96660	hypothetical prote
22	153	3.6	1250	2 T00454	hypothetical prote
23	152.5	3.6	202	2 F88469	protein F2K11.14 [
24	151.5	3.5	1350	2 T42697	protein C28H8.9 [i
25	151.5	3.5	1722	1 I78879	hypothetical prote
26	150.5	3.5	728	2 S57142	retinoblastoma bin
27	148	3.5	312	2 G84472	hypothetical prote
28	148	3.5	530	2 I38558	hypothetical prote
29	146	3.4	429	2 C84640	Mi-2 autoantigen 2 similar to mammali

30	146	3.4	609	2 A43906	nuclear phosphopro
31	146	3.4	1479	2 T17401	transcription regu
32	143.5	3.4	1257	2 T01020	hypothetical prote
33	142.5	3.3	796	2 T08555	pathogenesis-relat
34	139	3.3	449	2 T12495	hypothetical prote
35	138	3.2	1560	2 I54361	SMCX protein - hum
36	137	3.2	1146	2 H96796	hypothetical prote
37	136.5	3.2	1033	2 I48775	Smcx protein (esca
38	135.5	3.2	381	2 S25848	polyubiquitin 5 -
39	135.5	3.2	381	2 JCS489	polyubiquitin 5 -
40	135	3.2	379	2 S55245	polyubiquitin 5 -
41	135	3.2	1871	2 D96796	probable heat shoc
42	134.5	3.1	1088	2 T14917	homeotic protein P
43	134.5	3.1	1378	2 G88637	protein F53H1.4 [i
44	133.5	3.1	229	1 A56582	polyubiquitin - Eu
45	133.5	3.1	846	2 T25234	hypothetical prote

ALIGNMENTS

RESULT 1

E96612

Probable transcription factor F12K22.14 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C;Accession: E96612

R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonzo, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: E96612

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-641 <STO>

A;Cross-references: UNIPROT:Q9FVS3; GB:AE005173; NID:gl1079528; PIDN:AAG29238.1; GSPDB:GN

C;Genetics:

A;Gene: F12K22.14

A;Map position: 1

C;Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

Query Match 14.0%; Score 600; DB 2; Length 641;

Best Local Similarity 26.5%; Pred. No. 8.8e-33;

Matches 172; Conservative 68; Mismatches 160; Indels 248; Gaps 19;

QY 318 CHLCGRQDPDKQMCDECMAFHYCDPLPSSVPSDEWYCPEDCNDASEVVLG--- 374

Db 15 CMRCKSNPPPEESLTGCTCTPTWVHVSCLSSPPKTLASTLQWHPDCSGIDPLPVSGGAT 74

QY 375 -----ERLRKSK----- 381

Db 75 GFESAGSDLVAAIRAIDEADESLSTEKAKMRQLLSGKGVDEEDDEEKKKKGKGNPML 134

QY 382 ----- 381

Db 135 DVLSALGDNLMCSFCMQLPFPVTKPCGHNACLKFEKWMGQKRTCGKRSIIPEKMAK 194

QY 382 -----KNAKMASATSS-----SQRD-----WGKMACVGR 407

Db 195 NPRINSLVAIRLAKVKSAAITTSKVFIISNQDRDPDKAFTTBRAKTKGANAAG-- 252

QY 408 KECTIVPSNHYGPIP-----GIPVGTWRRPVQVSESGVHRPHVAGIHRGNDGSY 458

Db 253 KIYVTPDHRGPIPAENDPVRNQGLLVGESWEDLECRQWGAHPHVAGIAGQSTYGAQ 312

QY 459 SILVAGGYEDVDHGNFFTYTGGSGRDLGSKNRT-AEQSCDQKLTNTNRLALNCFAPIN 517

Db 313 SVALSGGYKDDDEHGEFLYTGSGGRLSGNKRNTKEQSFQKPEKSNAAKLJLSC----- 367  
QY 518 DOEGAEAKDWSGKPKVVRNVKGGKSKYAPAGNRYDGIYKVKYKPEKGGKGLVWR 577  
Db 368 -----KLGIPVRVRSHK-EKRSAYAPEGVRYDGYRIEKW---RKVGQVCR 413  
QY 578 YLLRRDDDEPGFWTK-EGKDRIKKGLTMOYEGYLEALANRERKENKREBESEEOBGG 636  
Db 414 YLFVRCDNEPAPWTSDENGDRPPI---PNIPE--LNMATDLFERKETPESWDFDE----- 463  
QY 637 PASPRTKGKWKRSAGGPRASGPRTSKTKVPEYSLTAQOSSLIREDKSNAKLWNE 696  
Db 464 -----GEGCKWM-----KPPASKKS-----VNVLAPEERKNLR----- 493  
QY 697 VLASLKORPASGPFQFLSKVEETFOCICQOELVERPITTVCOHNVCKDCLDRSF----- 752  
Db 494 -----KAIKAASNTWRALLKEFKCQICQOVLTLPTVTPCAHNFCKACLEAKFAGKT 546  
QY 753 -----RAQVFCPCRYDLGRSYA-MQVNOPLQTVNL 785  
Db 547 LVRRSTGGRTLRSRKNVLCPCPTDISDFLQNPQVNRVAEVIK 594

RESULT 2  
H96684  
probable RING zinc finger protein F15E12.8 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
A:Accession: H96684  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzborg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: H96684  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-622 <STO>  
A:Cross-references: UNIPROT:Q9C8E1; GB:AE005173; NID:g11038468; PIDN:AAG27747.1; GSPDB:G  
C:Genetics:  
A:Gene: F15E12.8  
A:Map position: 1  
C:Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

Query Match 13.0%; Score 556; DB 2; Length 622;  
Best Local Similarity 26.2%; Pred. No. 8.1e-30;  
Matches 165; Conservative 60; Mismatches 171; Indels 234; Gaps 20;  
QY 318 CHLCGRQDPDKQLMCDCECDMAFHLYCLDPPPLSSVSEDEWYCPCCR----- 364  
Db 15 CMRCQVTPPSEETLTGCTCTVFWHVSCLLPE-SLASSTGDWECPCDGVVVPVSAAPGTGI 73  
QY 365 ---NDASEVVLAGE---LRESKNAKMASATSSSQD----- 396  
Db 74 SGFESSGVLVAIRAIAQADVTILTEAEKAKKQRLMSGGGDGDVDDDEKKLEIFCSICI 133  
QY 397 -----WGKG-----MAC-----VG 405  
Db 134 QLPERPVTTPCGHNFCLKCFEKWAVGGKLTMCICRSKIPRHVAKNPRINLALVSAIRLA 193  
QY 406 RTKECT-----IVPSNHYGPIC- 422  
Db 194 NVTKCSGEATAAKVHHIIRNQDRPKAFTTERAVTKGANAASGKFFVILPRDHFQPIA 253  
QY 423 -----GIPVGTWVRFRVQVSESGVHRPHVAGIHRGNDGSYSLVLAGGYEDDGDHGN 474

Db 254 ANDVTRNQGVLCGESWEDRQECRQWVHFPHVAGIQAQAAVGAQSVASLGGYDDEHGE 313  
QY 475 FFTYTGSGGRDLGNKKTAB-QSCDQKLTNTNRALALNCFAPINDQGAEKAKWRSKGPV 533  
Db 314 WFLYTGSGRDLGNKRNKVNKIQSSDQAFKNMNEALRLSC-----KMGTPV 358  
QY 534 RVVENVKGGKSKYAPAGNRYDGIYKVKYKPEKGGSGF-LVMRYLLRRDDDEPGFWTK 592  
Db 359 RVRSWK-EKRSAYAPEGVRYDGYRIEKWCVNGVQGLHKMCRYLTVRCNDEPAPWTS 417  
QY 593 -EGKDRIKKGLTMOYEGYLEALANRERKENKREBESEEOBGGFASPRTKGKWKRS 651  
Db 418 DEHGRPRPL---PDVPE--LENATDLFVRKESPSWGFDEAE-----GRWKWKMS 463  
QY 652 AGGPRASGPRTSKTKVPEYSLTAQOSSLIREDKSNAKLWNEVLASLKDRPASGPF 711  
Db 464 P---VSRMALDTEERKKNKRAKGNNAKARLLKE----- 495  
QY 712 QLFLSKVEETFOCICQOELVERPITTVCOHNVCKDCLDRSF-----R 753  
Db 496 -----FSCQICRKVLSLPTVTPCAHNFCKACLEAKFAGITQLRDRSNGVRKLR 545  
QY 754 AQVFCPCRYDLGRSYA-MQVNOPLQTVL 782  
Db 546 KNIMTCCPCTTDLSEFLQNPQVNRVEMBEII 575

RESULT 3  
A96685  
probable RING zinc finger protein F15E12.5 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
A:Accession: A96685  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzborg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: A96685  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-598 <STO>  
A:Cross-references: UNIPROT:Q9C8E0; GB:AE005173; NID:g11038479; PIDN:AAG27758.1; GSPDB:G  
C:Genetics:  
A:Gene: F15E12.5  
A:Map position: 1  
C:Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

Query Match 12.5%; Score 533.5; DB 2; Length 598;  
Best Local Similarity 26.8%; Pred. No. 2.6e-28;  
Matches 162; Conservative 54; Mismatches 180; Indels 209; Gaps 19;  
QY 318 CHLCGRQDPDKQLMCDCECDMAFHLYCLDPPPLSSVSEDEWYCPCCR----- 365  
Db 15 CMRCQVTPPSEETLTGCTCTVFWHVSCLLPE-SLASSTGDWECPCDGVVVPVSAAPGTGI 73  
QY 366 -----DASEVVLAGEIRRESKNAKMASATSSSQD----- 396  
Db 74 SGFESSGVLVAIRAIAQADVTILTEAEKAKKQRLMSGGGDGDVDDDEKKLEIFCSICI 133  
QY 397 -----WGKG-----MACVGR----- 406  
Db 134 QLPERPVTTPCGHNFCLKCFEKWAVGGKLTMCICRSKIPRHVAKNPRINLALVSAIRLA 193  
QY 407 --TK---ECTIVPSNHY-----GPIPIGVGTWVRFRVQVSES 439  
Db 194 NVTKCSGEATAAKVHHIIRNQDRPKAFTTERAVTKGANAASGLVGVSWEDRQECRQW 253



QY 440 GVHREPHVAGIHRGSRNDGSYSVLVAGGYDDVDHGNNFFTYTGGGRDLSGNKRTAE--QSCD 498  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 254 GVHFPHVAGIAGQAAGVGAQSVASGGYDDEDHGEWFLYTSGGRDLSGNKRVRNKIQSD 313  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 499 QKLITWNPALINCFAPINDQEGABAKWRSGKPVVVRYNYKGKNKYAPAENGRYDGI 558  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 314 QAFKMNEALRLSC-----KMGYPVVRVRSWK-EKRSAYAPAEGRYVDGV 357  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 559 YKVVKYWEPEKSGSF-LVMRYLLRRDDDPGPWTYK-EGKDRIKKLGLTMQYPEGYLEALA 616  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 358 YRIEKCWNVGQGLHKMCRVILFRCONEPAPWTTSDEHGDPRFL---PDVPE--LENAT 412  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 617 NREREKENSEEBEQEGGFASFPTGKGNKRWKSAGGGPSRAGSPRRTSKTKTVEPYSL 676  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 413 DLFVRKESPSGFEAE-----GEWKMKPPP---VSFMALDTEERKANRAKNGN 460  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 677 TAQSSLIREDKSNAKLNWEVIASIKORPASGPSFQLFKSVERTFCICQCLVFRRPIT 736  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 461 NAMKARLUKE-----FSQCICRKVLSPVT 485  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 737 TVCOHNVCCKCLDRSF-----RAQVFCPCACRYDLGRSYA-MQNQP 777  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 486 TPCAHNFCACLEAKFAGITOLDRSNGVRKLRAKNIMTCPCCTTLDFELQNPVNRE 545  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 778 LQTVL 782  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 546 MMEII 550  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
  
RESULT 4  
D96612  
hypothetical protein F12K22.15 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: D96612  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: D96612  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-650 <STO>  
A:Cross-references: UNIPROT:Q9FVS2; GB:AEO05173; NID:g11079520; PIDN:AAG29230.1; GSFD:  
C:Genetics:  
A:Gene: F12K22.15  
A:Map position: 1  
C:Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger  
  
Query Match 11.1%; Score 473; DB 2; Length 650;  
Best Local Similarity 23.6%; Pred.No.3.6e-24;  
Matches 155; Conservative 71; Mismatches 168; Indels 264; Gaps 21;  
  
QY 318 CHLCGRQDPDKQALMCDCDMAPHYICLPDPLSSVSEDEWYCERN----- 365  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 15 CMRCKSNPPPEESLTGTCTVPMHVSCLLSPPETLSATQWLCPDCSGETNPPLFVSGVAA 74  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 366 -----DASEVLAGERLR-----ESKNAK----- 385  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 75 GYGSGVSDLVAAIHSIEADETLSAEAEKAKKQQLSGKWVDEDEEEKKTSKGKPKID 134  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 386 ----- 385  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 135 VLSHFECFGCMQSLQKPVSVRVLFALMLVFWLESTPCGHNAACLFLKMWQGHRSCG 194  
||| ||| ||| : : : ||| : | : | : | : | : | : | : | : | : | : | : | : | :

QY	386	-----MASATSS-----SORDW	397
Db	195	TCRSVIPESMVTNPRINLSIVSAILARVSEKADARTSKVVHYVDNEDRDPKAFTERAK	254
QY	398	GKGMACVGRTEKCTIVPSNHYGPIP-----GIPVGTWRFVRVQVSESGVHRPHVAG	448
Db	255	KTGNANASSGKIFVTIPRDHFGPIPAENDPVRNOGLLVGSEWKGRLACRQWGAHPHVS	314
QY	449	IHGSRNPGSYSLVLAGGYEDVDHGNFTYTGSGGRDLISGNKRTAEQSCDQKLTNTNRAL	508
Db	315	IAGQASYGAQSVVLAGGYDDDEDHGEWFLYT-----RTNVTQAFQVFLNFNEAL	365
QY	509	ALNCFEAPINDOGEAEKDWKSGKPVVRNVNKGKSNKYAPAEQ-NRYDGIYKVVKWPE	567
Db	366	RLSC-----KLGYAPVRVVRSTK-DKRSYPAQGGLLRYDGVYRIEKCW--	407
QY	568	GKSGGFLVWRYLLRRDDDEPGWTK-EGKDRIKKLGLTMQYPEGYLEALANRERENSK	626
Db	408	RIVGIGQKCPFLVRCONEPAPMTSDHGGORPRL-----PNVPE--LNMAITDLPERKESPS	461
QY	627	REEBEOQEGFPASPRTGKWKGRKSKAGGSPRAGSPRRTSKTKTKVEPYSYLTAAQOSSLIRE	686
Db	462	WDFDE-----GEDRW-----RWMKPPPPASKKAVKNVL---	488
QY	687	DKSNAKLWNEVLASLKDORPASGSPFQFLFSKVBTFCICQBELVFRPITTVCOHNVC	746
Db	489	DPEERKLLREAI-----KSANP-NTMARLLKEFKQCICQKVMNTPVTTCAHNFC	540
QY	747	CLD-----RSFRAQ--VFSCPRCYDLGRSYA-MQVNPLOTVLNLQ	785
Db	541	CLESKFAGTALVRGSGGRKURSKQSMKPCPCPTDIAEFVQNPQVNRVAVIEKL	598
RESULT 5			
T01825			
hypoetical protein T15F16.7 - Arabidopsis thaliana			
C/Species: Arabidopsis thaliana (mouse-ear cress)			
C/Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004			
C/Accession: T01825			
R/Antoniou, B.; Le, T.			
submitted to the EMBL Data Library, August 1998			
A/Description: The sequence of A. thaliana T15F16.			
A/Reference number: Z14443			
A/Accession: T01825			
A/Status: translated from GB/EMBL/DDSB			
A/Molecule type: DNA			
A/Residues: 1-461 <ANT>			
A/Cross-references: UNIPROT:O81463; EMBL:AF076275; NID:g3293582; PID:g3377817			
A/Experimental source: cultivar Columbia			
C/Genetics:			
A/Map position: 4			
A/Introns: 121/3; 296/2; 329/2; 362/3; 380/2; 414/3			
A/Note: T15F16.7			
C/Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger F;105-153/Domain: RING finger homology <RRR>			
Query Match			
Best Local Similarity 30.2%; Pred. No. 3e-21;			
Matches 136; Conservative 54; Mismatches 146; Indels 114; Gaps 21;			
QY	263	DSGLNDCRIIFVDEVPKIERPGEQSPWVDNPMRRKSGPSCKHCKDDVNRLLCRVCA---CH	319
Db	58	DSMTD-----ADETKKRKLISLGCEADE--NNKSDGEIASLNDGVDAFTAICDLNCS	110
QY	320	LCGGQDPDKQLMCECDVAFHYICLDPPLSSVPSEDEW-----YCPECRNDASEVILA	373
Db	111	LCNQLPDRPVTILCG-----HNFLC-----KGFDKWIDQGNQICATCRSTIPDKMAA	157
QY	374	GERLRES-----KKNAKMASATS-----SSQSDWKGMAVGRTEKCTI-----	412
Db	158	NFRVNSLSVTRYVKVAKTAGVFANFPPTFSNQD-----GPNAFRTKRAKIGEENARI	214
QY	413	---VPSNHYGPIP-----GIPVGTWRFVRVQVSESGVHRPHVAGIHGRSNDGSYSL	460

Db 215 YVTVFDFHFGFIPAEHDFVRNQVLGVGESWENRVECKQWGHLPVHVSCTAGQEDYGAQSV 274  
QY 461 VLAGGYDDVDHGNFFYTG-SGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQ 519  
Db 275 VISGGYKDDDEHGWFLYTGSRGRHPANE-----DQEFEDINEALRVSC----- 319  
QY 520 EGAEAKDWRSGKPVVRVNRVNGKGNKYAPAEGRNRYDGIYKVYKWPKEGSGFLVWRYL 579  
Db 320 -----EMGYPVVRVRSYK-DRYSAYPAREGVYDGVYRIEKCW---RKARFPVCRYL 367  
QY 580 LRRDDDEPGPW-TKEGDKRIKGLTQWYPEGYLEALANRREKENSKEEEEOQEGGFA 638  
Db 368 FVRCDNEPAPNWSDEGDRPRPL---FNPIE--LETASDLFERKESPSWDFDEAE----- 417  
QY 639 SPRTGKGKWKKSAGGSPRAGSPRRTSKK 668  
Db 418 ----GRWRWK-----PPPANHEQRERMK 437

RESULT 6  
T00949  
conserved hypothetical protein T3F12.10 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T00949  
R:Gnoj, L.; Huang, E.N.; Habermann, K.; Hameed, A.; Hasegawa, A.; Jensen, K.; Schutz, K.  
R.; McCombie, W.R.  
submitted to the EMBL Data Library, October 1997  
A:Description: Arabidopsis thaliana BAC T3F12 from chromosome IV.  
A:Reference number: Z14210  
A:Accession: T00949  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-432 <GNO>  
A:Cross-references: UNIPROT:O22280; EMBL:AC002983; NID:g2443899; PID:g2565009  
C:Genetics:  
A:Map position: 4  
A:Introns: 121/3; 296/2; 329/3; 351/2; 385/3  
A:Note: T3F12.10  
C:Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger  
F105-153/Domain: RING finger homology <RRN>

Query Match 8.3%; Score 353.5; DB 2; Length 432;  
Best Local Similarity 26.9%; Pred. No. 2.6e-16;  
Matches 121; Conservative 49; Mismatches 137; Indels 143; Gaps 19;

QY 263 DDSLNDCRIFVDEVFKEIRPCEGSPWVDNPMRRKSGPSCKHKDDVNRLCRVCA---CH 319  
Db 58 DESMTD-----ADETKKRKRIILSGDCEAD--NNKSDGELASINDGVDAFTAICEDLNCS 110  
QY 320 LGGGRQDPDKQMLCDCECDMAFIYCLDPLPSSVPSSEW-----YCPCEKNDASEVIL 373  
Db 111 LCNQLPDRPVTILCG-----HNFCL-----KCFDKWIDQGNQICATCRSTIPDKMAA 157  
QY 374 GRLRES-----KKNAKMASATS-----SSQPDWGMGMAVCGTKECTI----- 412  
Db 158 NRVNSSLVSVIRYVKVAKTAGVTANFPFPFTSNQD---GPNAPFTKRAKIGENNAARI 214  
QY 413 ---VPSNHYGPIP-----GIPVGTWRRFVQVSESGVHRPHVAGIHGRSNDGSLSL 460  
Db 215 YVTVFDFHFGFIPAEHDFVRNQVLGVGESWENRVECKQWGHLPVHVSCTAGQEDYGAQSV 274  
QY 461 VLAGGYDDVDHGNFFYTG-SGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQ 520  
Db 275 VISGGYKDDDEHGWFLYTGSRGRHPANE-----DQEFEDINEALRVSC----- 295  
QY 521 GAEAKDWRSGKPVVRVNRVNGKGNKYAPAEGRNRYDGIYKVYKWPKEGSGFLVWRYL 579  
Db 296 -----RSYK-----DRYSAYAPAREGVYDGVYRIEKCW---RKARFPVCRYL 338  
QY 580 LRRDDDEPGPW-TKEGDKRIKGLTQWYPEGYLEALANRREKENSKEEEEOQEGGFA 638

Db 339 FVRCDNEPAPNWSDEGDRPRPL---FNPIE--LETASDLFERKESPSWDFDEAE----- 388  
QY 639 SPRTGKGKWKKSAGGSPRAGSPRRTSKK 668  
Db 389 ----GRWRWK-----PPPANHEQRERMK 408

RESULT 7  
C75384  
conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: C75384  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: C75384  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-299 <WHI>  
A:Cross-references: UNIPROT:Q9RU61; GB:AE001997; GB:AE000513; NID:g6459292; PIDN:AAF1109  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DRI533  
A:Map position: 1

Query Match 7.3%; Score 311.5; DB 2; Length 299;  
Best Local Similarity 38.2%; Pred. No. 1.1e-13;  
Matches 65; Conservative 34; Mismatches 50; Indels 21; Gaps 3;

QY 417 HYGPIPIPVGTWRRFVRVQVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYDDVDHGNFF 476  
Db 4 HFGAVGVGVPGMAFVNROELRDAGVHLPTQAGISGASEGADSVLSGSGYEDDRDEGDI 63  
QY 477 TYTSGGRDLGSGNKRTEAQSCDQKLTNTNRALALNCFAPINDQGAEAERKMRSGKPVRV 536  
Db 64 LYTGGGRD-----PLTGHQVFPQQLVRGNLALAIS-----HRDGLPLRV 104  
QY 537 RNVGKGNKYAPAEGRNRYDGIYKVYKWPKEGSGFLVWRYLRRDDDE 586  
Db 105 RGH--HSSQFSPQSGYVAGLYRVDDHWRVGRSGFLIWRPLRLTLENQ 152

RESULT 8  
D84765  
similar to mammalian MHC III region protein G9a [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: D84765  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: D84765  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-794 <STO>  
A:Cross-references: UNIPROT:O82175; GB:AE002093; NID:g3668088; PIDN:AAC61820.1; GSPDB:GN  
C:Genetics:  
A:Gene: At2g35160  
A:Map position: 2

Query Match 7.0%; Score 299.5; DB 2; Length 794;  
Best Local Similarity 23.7%; Pred. No. 2.5e-12;  
Matches 136; Conservative 75; Mismatches 231; Indels 133; Gaps 20;

QY 48 RQKQMEDGHTLPDYEVRLNDTTLQLLRQSLVLPKSTKRDSELSLSDTSGCCLQSGSESDKS 107

Db 65 RSRMRQKEFTVETETR-----NVSDV-----CVLSQADVE 96  
QY 108 STGEEAAETDSRPADDEMDDETLGLYKNVEYVDARTNGAFEAQVVRVTRKAPSRD 167  
Db 97 LIPGEIWAERDSFKS-----VDCNDMSVGLTEGAESLGVNNQEPMKD 138  
QY 168 EPC-SSTSRLPALEDVIVHKYDDYPENGVVQMSR-----DVRARARTIKWQDLEVGQ 221  
Db 139 RNPENTISEQNWVE---VHPSPISLPEDMGMSVCRKSITGKELHGRITISVGRDLSPNM 195  
QY 222 VMLNVPNDPKRPFWDABISKRRTARELYANVVL-----GDSLNDCKRIIFVD- 275  
Db 196 GSKFSKNGKTAKR-----SISVEEB-----NLVLEKSDSGHGLSPPEVLELEK 239  
QY 276 -EYFKIIPRPEGSPMVDNPMRRKSPCKHCKDDVNLRCVACHLCGGQDDPKQLMCD 334  
Db 240 SEVMIITDKGVWMPSPVKPEKNG-----DYGE-----GSMRKNSRVALD 281  
QY 335 ECDMAFIYCLDPLSLSPSEDEWYQPCRNDASEVVLAGERL-RESKNAKVASATSSS 393  
Db 282 KKLASKFLSNGGLPS-----CSSGDSARYKVKTETWRLPHETCKKIMQBEAEARPR 333  
QY 394 ORDWKGMACVGRKTECTIVPSNHV-----GPIPIPVGTMMFRVQVSESGVHRPHVA 447  
Db 334 KRDCGNPKVVCESAKILSKGKLYSCTQIIIGTPGVGEVDFQYRMELNLGHRPSQS 393  
QY 448 GIHGRSNDG---SYSLVLAGGYEDVDHGNFTYTGSGGRDLSGNKRTABQSCDKLTN 503  
Db 394 GIDYMKDDGELVATISVSGGYNDVLDNSDVLITYTGGGN--VGKKKNNEPRKDDQLVT 451  
QY 504 TNRALALNCFAPINDQGEAKDWRSGKPVVRVNRVNGKNSKYAPAEGRNRYDGIYKVKV 563  
Db 452 GNALAKNS-----INKK-----NPRVIRGINKNTYLQSSVVAKNVYDGLYLVEE 496  
QY 564 YWPEKKGSGFLVWRYLLRRDDDFG-PWTKEGKDR 597  
Db 497 YWBEETSGHKLVEFKLRRIPGQPELPWKEVAKS 531

## RESULT 9

C84616  
similar to mammalian MHC III region protein G9a [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: C84616  
R.;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: C84616  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-788 <STO>  
A;Cross-references: GB:AE002093; NID:g4314371; PIDN:AAD15582.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g22740  
A;Map position: 2

Query Match 6.2%; Score 265.5; DB 2; Length 788;  
Best Local Similarity 21.9%; Pred. No. 5.1e-10;  
Matches 143; Conservative 94; Mismatches 218; Indels 199; Gaps 27;

QY 16 VDSLSRLTKVEELRRKIQLFHVPEGL-QRLFYRGKQMEDGHTLFDYVRLNDTIQLLVR 74  
Db 37 VCKLDRMSGKLFKRRKVFVAVRDPFGGSGRAMEVKIACENGNNVEDVKV-----VSSLVK 91  
QY 75 QSLVLPHSTKERSDSELSDDT-----SCCLL---GQSESDKSTTGEAAETDSRP 121  
Db 92 EE-----ESLQORDASENSVDIRMAEPVEVQPLRICLPGGVDVVRDLSTVATGDCSNEQIV 147  
QY 122 ADEDMWDE--TELGLYKNVEYVDARTNNGAFEAQVVRVTRKAPSRDEPCSTSRPALE 179

Db 148 AGSSVSSSGTENIVRDIVVYADESSLSGMDNLDOTQPLEIEMSDVAVAKPLVAGRKAK 207  
QY 180 EDVLYHKYDDYPENGVVQMSRDRARARTIKWQDLEVGQVVMNLVNPDPKRGFWY 239  
Db 208 KGTAACH-----SSLKVVSREFG----- 224  
QY 240 DAETSRKRETRTARELYANVVLGDSLNDCKRIIFVDVFFKIEREGEGSPMVDNPMRRKSG 299  
Db 225 --EGSRKKKSK--KNLYWRDRRESLDSPEQLRILGVG-----TSSGS-----SSG 264  
QY 300 PSCKHCKDDVNLRCVACHLCGGQDDPKQLMCDCEMAFIYCLDPLSLSPSEDEWY 359  
Db 265 DSRNKVKETLRL-----FHGYC-----RKILQDEE-- 290  
QY 360 CPECRNDASEVVLAGERLRESKNAKVASATSSORDWKGKMACVGRKTECTIVPSN-- 416  
Db 291 -----AKPEDORKKGKGLRI--DFEASTILKRNKGF 319  
QY 417 -----H-YGPIPIPVGTMMFRVQVSESGVHRPHVAGI-----HGRSNDGSYSLVLAGGY 466  
Db 320 LNSGVHILGEVPGVGEVDFQYRMELNLGIHKPSQAGIDYMKYGAQKAVAT-SIVASGGY 378  
QY 467 EDDVDHGNFTYTGSGGRDLSGNKRTAB--QSCDKLTNTNRALALNCFAPINDQGEAEA 524  
Db 379 DDHLNDSVLITYTGGGNVNMVKKGEELKEPEDQKLTITGNLALATS-----IEKQ----- 429  
QY 525 KDWRSKPVVRVNRVNGKNSKYAPAEGRN--YDGIYKVKVYKPEKSGFLVWRYLLRR 582  
Db 430 -----TPVERVIR--GKHKSTHDKSGGNYVDGLYLVKEKYQQVQVGHGSHGMVFKFQLRR 480  
QY 583 DDDEPG-PWTKEGKDRIKKGLTQMYPEGY--LEALANREREREKENSKEEEBQQ 633  
Db 481 IPGQPELSVWEVKKSKS-----YREGICKLDSBGKEQSPISAVNETDDEK 527

## RESULT 10

T06648  
hypotheical protein T6G15.10 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: T06648  
R.;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, April 1999  
A;Reference number: Z15791  
A;Accession: T06648  
A;Molecule type: DNA  
A;Residues: 1-650 <BEV>  
A;Cross-references: UNIPROT:Q9T0G7; EMBL:AL049656; GSPDB:GN00062; ATSP:T6G15.10  
A;Experimental source: cultivar Columbia; BAC clone T6G15  
C;Genetics:  
A;Gene: ATSP:T6G15.10  
A;Map position: 4

Query Match 5.4%; Score 229.5; DB 2; Length 650;  
Best Local Similarity 28.8%; Pred. No. 1.1e-07;  
Matches 74; Conservative 39; Mismatches 83; Indels 61; Gaps 11;

QY 372 LAGERLRESKNAKVASATSSORDWKGKMACVGRKTECTIVPSNHVGTPIGIPVGTMMWR 431  
Db 172 VGRRARADCK-----AGKAGSMWRDC---MLWNRRDKRIV-----GSPGVQVGDIF 217  
QY 432 FRVQVSESGVHRPHVAGTH---GRSNDG---SYSLVLAGGYEDVDHGNFTYTGSGGR 484  
Db 218 FRFELCVMLGHHPQSGIDFLTGLSSNGEPTATSVIVSGYEDDDDDQGDVINYTGQGGQ 277  
QY 485 DLSGNKRTABQSCDKLTNTNRALALNCFAPINDQGEAKDWRSGKPVVRVNRVNGKSN 544  
Db 278 D-----RLGROAHOHLRLEGGNLAERSMY-----YGIEVRVIRGL----- 312  
QY 545 SKYAPAEGRN---YDGIYKVKVYKPEKSGFLVWRYLLRRDDDFGPGWTKEGKDRIKKL 601  
Db 313 -KYENEVSRTVYDGLFRIVDSWFDVGKSGGVFKYRLRIEQQ-----AEMGSSVLKF 366

QY 602 GLTMQ-----YPEGYL 612  
DB 367 ARTLTKNPLSVRPRGI 383

RESULT 11  
F94743  
similar to mammalian MHC III region protein g9a [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: F94743  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: F94743  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-651 <STO>  
A:CROSS-references: UNIPROT:O22781; GB:AE002093; NID:g2459412; PIDN:AAB80647.1; GSPDB:GN  
C:Genetics:  
A:Gene: At2g33290  
A:Map position: 2

Query Match 5.3%; Score 226; DB 2; Length 651;  
Best Local Similarity 26.4%; Pred. No. 1.9e-07;  
Matches 85; Conservative 55; Mismatches 106; Indels 76; Gaps 17;

QY 345 LDPP-----LSSVPSDEWYPCERNDASEVVLG---ERLRKSKNAKMA----- 387  
DB 108 LEPPPGFKDNRVSTVVSFKPERP---RELARIAIIGHEQRKELQVWKRTMTVESLRI 164  
QY 388 --SATSSQRDWMGK-----MACVGRTEKCTTVPSNH-YGPIPGIPVGMFRFRVQV 436  
DB 165 HLMAESKNHVLGQRRRSDMAAIVMDRGLWLNVDYKHIVGPTGVVEGDIFFYRMEL 224  
QY 437 SESGVRHPRVAGIH-----GRSNDG---SYSVLVAGYEDVDVHGFFYTGSGGRDLGN 489  
DB 225 CVLGLHGTOAGIDCLTBSATGPFIATSIYVSGYEDDEDTGDLVYTGHGQD---- 280  
QY 490 KRTAQSCD-QKLTNTNRALNCFAPINDQGAERAKDWRSGKPKRVVRNVKGGKNSKYA 548  
DB 281 --HQHKQDNQLVCGNLG-----ERSMHYGLVVRVIRGI-----KVE 317  
QY 549 PAEGNR---YDGIYKVYKVPKSGKGLVWRVYLLRRDDDP--GPWTKGDKRIKKLGL 603  
DB 318 NISSKVVYVDGLYKIVDWFVAVGKSGFGVFRVLRIBGQPMGSAVMRFAQTLRNKP- 376  
QY 604 TMQYPEGYLE-ALANRERKEN 624  
DB 377 SMVREPTGYVSFDLSNK---KEN 395

RESULT 12  
F96756  
hypothetical protein F3N23.30 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: F96756  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F96756  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-869 <STO>  
A:CROSS-references: GB:AE005173; NID:g5903099; PIDN:AAD55657.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F3N23.30  
A:Map position: 1

Query Match 4.6%; Score 194.5; DB 2; Length 669;  
Best Local Similarity 29.5%; Pred. No. 2.6e-05;  
Matches 67; Conservative 31; Mismatches 92; Indels 37; Gaps 8;

QY 419 GPIPGIPVGTWVRFRVQVSESGVHRPHVAGIH-----GRSNDG---SYSVLVAGYEDVD 471  
DB 208 GTVPGIEVGDIFFSRIENCLVGLHMQTVAGIDIYISKAGSDEESLATSIVSSGRYEQAQ 267  
QY 472 HGNFTYTGSGRDLGSKNRTABQSCDKLTNTNRALNCFAPINDQGAERAKDWRSGK 531  
DB 268 DPESLIYSGQG-----GNADKNRQASDQKLRGNLAL-----ENSLRKN 307  
QY 532 PVRVVRNVKGGKNSKYAPAEGNRVDGIYKVYKVPKSGKGLVWRVYLLRRDDDP--G 588  
DB 308 GYRVVR---GBEDAASKTKGIYIYDGLYSISSEWVEKSGKSGNTFKYKLVROPQGPAPG 364  
QY 589 PWTKEGDKRIKKLGLTMOYPEGYLEALANRERERKENSXREEEQQEG 635  
DB 365 FWKSVQK---WKEGLTTR-PGLILPDLTSGAESKPSVLVNDVDEDKG 407

RESULT 13  
G86312  
hypothetical protein F2H15.1 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
C:Accession: G86312  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G86312  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-954 <STO>  
A:CROSS-references: GB:AE005172; NID:g965056; PIDN:AAF97258.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1

Query Match 4.5%; Score 192.5; DB 2; Length 954;  
Best Local Similarity 30.6%; Pred. No. 5.6e-05;  
Matches 55; Conservative 28; Mismatches 58; Indels 39; Gaps 6;

QY 419 GPIPGIPVGTWVRFRVQVSESGVHRPHVAGI-----HGRSNDGYSVLVAGYEDVD 471  
DB 227 GAVPGIHVGDIFFYWGEMCLVGLHKSNGYGGIDFFTAASAVEGHAMCVVTAGQYDGETE 286  
QY 472 HGNFTYTGSGRDLGSKNRTABQSCDKLTNTNRALNCFAPINDQGAERAKDWRSGK 531  
DB 287 GLDTLIYSGQGTVDYGNAR-----DQMKGNLAL-----EASVSK---GN 325  
QY 532 PVRVVRNVKGGKNSKYAPAEGNR---YDGIYKVYKVPKSGKGLVWRVYLLRRDDDP 587  
DB 326 DVRVVRGV-----IHPHENNQKIYIDGMVLSKFWTVTKSGKGFKEFRFLVRKNQF 378

RESULT 14

T03455  
ALR protein - human  
C:Species: Homo sapiens (man)  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C/Accession: T03455  
R:Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano, A.; Title: Structure and expression pattern of human ALR, a novel gene with strong homology Oncogene 15, 549-560, 1997  
A:Reference number: Z14954; MUID:97388474; PMID:9247308  
A:Accession: T03455  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-4957 <PRA>  
A/Cross-references: UNIPROT:O14686; EMBL:AF010404; NID:G2358286; PIDN:AACS1735.1; PID:G2358286  
C:Genetics:  
A:Gene: ALR  
A:Map position: 12  
C:Superfamily: acute lymphoblastic leukemia protein, ALR type  
C:Keywords: alternative splicing

Query Match 4.2%; Score 177.5; DB 2; Length 4957;  
Best Local Similarity 20.2%; Pred. No. 0.0048;  
Matches 156; Conservative 88; Mismatches 309; Indels 221; Gaps 34;

QY 88 SELSDTDSGCCGCGSE-----SDKSSTHGEAAATDSRPA-DE-----DMWDETE 131  
Db 629 SNLSQGDASASFPGSEPLGSPDEGGGSLMELGVSTDVSPARDEGSLRLCTDSLPETD 688  
QY 132 LGLYKNEYVDARTNMGAFQAQVVRVTRKAPSRDEPCSTSRPALEEDVIVHKYDDY 191  
Db 689 DSL-----LCDAGTAISGKAEKGR-RRSSPARSR-----IKQGRSSFF 728  
QY 192 PENGVMQNSRDVRARARTIIKQDLEVGQVMVNLNPNPKERGFWDYAEISKRRTTRT 251  
Db 729 PGRRRPGGAHGRGRARLKTASSIETLVAD-----IDSSPSKEEED 776  
QY 252 ARELYANVLGDDSLNDCRIIFVDEVKI-----ERPGEQSPMDVNMRRKSGSCXCKD 307  
Db 777 DDTMONTVLF-----NTDKFVLMQDMCVCGSGRGAEGHLLACSCQSCYHYCVNSKI 833  
QY 308 DVNRL-----CRVC-ACHLCGGQDDPKQLMCDCECMAPHYICLDPLPLSSVPS----- 354  
Db 834 TKVMLLKGRWCVEICVCEVGQASDFSRLLDCDDISYHTYICLDPLPLTPVKGWCKW 893  
QY 355 -----EDEWY-----CPCRND-ASEVVLAGERLRESKNAK 385  
Db 894 CVSCMQCGAASPGFHCWQNSYTHCGPCASLVTCPICHAPYVEEDLLIQCRHCEWMHAG 953  
QY 386 MASATSSSQDDWG--KGMACVGRTECTIVPSNHGPIGPIVGT-----MWRFR-VQV 436  
Db 954 CBSLFTDDVDHAPDEGDCVS-CQPVYVVPVAPVAPPVPMKVKEPEQYFRFEGVWL 1012  
QY 437 SSSG-----VHR-----PHVAGIHGRSNDGYSYSLVLAGYEDVDVHGNF 475  
Db 1013 TETGMALLNLMTSPHLKRRQRRLGLPGEAGLEGSEPSDALG-----PDDKDGDL 1065  
QY 476 FT---YTSGGGRDLNKGKTAQSCDQKLTNTNRLALNCFAPIN-----DOGEAKDW 527  
Db 1066 DTDELLKGGG-----VEHMECEIKLE-----GPVSPDVPFGKEETESKK 1106  
QY 528 RSGKQVR-----VVRNVKGGKNSKAP-AEGRNRYDGIYKVKYWEKSGFLVWRYLL 580  
Db 1107 RKRKPYRPGIGGFMVQRKSHTRTKGPAQAQAEVLSDGQDPDEVIPADLPAGEAVEQSL- 1165  
QY 581 RRDDEPGPWTKEGDKRIKGLTQYPEGYLEALANRE----- 619  
Db 1166 -AEGDEKKKQRRGRKRSKLEG---FPAYLQEAFFGKELLDLSRKALFAVGVRPSFGL 1221  
QY 620 -----REKENSKR-----ESEEQOEGGFASPRCKGKWKRSAGGGS 657  
Db 1222 GTPKAKGDSGERKELPTSKQDGDGDIADERSGLEKADTPGPDGGVKASPVSPDPE 1281  
QY 658 RAGSPRTSKTKVFYSILTAQSSLIREDKSNAKIWNVEVLASLDRP-ASGSP 710

Db 1282 KPGTPGEGMLSSDLDRIS-TEELPKM--ESKDLQQLFXDVLGSEREQHLCGCTP 1332

RESULT 15  
T03454  
ALR protein - human  
C:Species: Homo sapiens (man)  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C/Accession: T03454  
R:Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano, A.; Title: Structure and expression pattern of human ALR, a novel gene with strong homology Oncogene 15, 549-560, 1997  
A:Reference number: Z14954; MUID:97388474; PMID:9247308  
A:Accession: T03454  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-5262 <PRA>  
A/Cross-references: UNIPROT:O14686; EMBL:AF010403; NID:G2358284; PIDN:AACS1734.1; PID:G2358284  
C:Genetics:  
A:Gene: ALR  
A:Map position: 12  
C:Superfamily: acute lymphoblastic leukemia protein, ALR type  
C:Keywords: alternative splicing

Query Match 4.2%; Score 177.5; DB 2; Length 5262;  
Best Local Similarity 20.2%; Pred. No. 0.0052;  
Matches 156; Conservative 88; Mismatches 309; Indels 221; Gaps 34;

QY 88 SELSDTDSGCCGCGSE-----SDKSSTHGEAAATDSRPA-DE-----DMWDETE 131  
Db 934 SNLSQGDASASFPGSEPLGSPDEGGGSLMELGVSTDVSPARDEGSLRLCTDSLPETD 993  
QY 132 LGLYKNEYVDARTNMGAFQAQVVRVTRKAPSRDEPCSTSRPALEEDVIVHKYDDY 191  
Db 994 DSL-----LCDAGTAISGKAEKGR-RRSSPARSR-----IKQGRSSFF 1033  
QY 192 PENGVMQNSRDVRARARTIIKQDLEVGQVMVNLNPNPKERGFWDYAEISKRRTTRT 251  
Db 1034 PGRRRPGGAHGRGRARLKTASSIETLVAD-----IDSSPSKEEED 1081  
QY 252 ARELYANVLGDDSLNDCRIIFVDEVKI-----ERPGEQSPMDVNMRRKSGSCXCKD 307  
Db 1082 DDTMONTVLF-----NTDKFVLMQDMCVCGSGRGAEGHLLACSCQSCYHYCVNSKI 1138  
QY 308 DVNRL-----CRVC-ACHLCGGQDDPKQLMCDCECMAPHYICLDPLPLSSVPS----- 354  
Db 1139 TKVMLLKGRWCVEICVCEVGQASDFSRLLDCDDISYHTYICLDPLPLTPVKGWCKW 1198  
QY 355 -----EDEWY-----CPCRND-ASEVVLAGERLRESKNAK 385  
Db 1199 CVSCMQCGAASPGFHCWQNSYTHCGPCASLVTCPICHAPYVEEDLLIQCRHCEWMHAG 1258  
QY 386 MASATSSSQDDWG--KGMACVGRTECTIVPSNHGPIGPIVGT-----MWRFR-VQV 436  
Db 1259 CBSLFTDDVDHAPDEGDCVS-CQPVYVVPVAPVAPPVPMKVKEPEQYFRFEGVWL 1317  
QY 437 SSSG-----VHR-----PHVAGIHGRSNDGYSYSLVLAGYEDVDVHGNF 475  
Db 1318 TETGMALLNLMTSPHLKRRQRRLGLPGEAGLEGSEPSDALG-----PDDKDGDL 1370  
QY 476 FT---YTSGGGRDLNKGKTAQSCDQKLTNTNRLALNCFAPIN-----DOGEAKDW 527  
Db 1371 DTDELLKGGG-----VEHMECEIKLE-----GPVSPDVPFGKEETESKK 1411  
QY 528 RSGKQVR-----VVRNVKGGKNSKAP-AEGRNRYDGIYKVKYWEKSGFLVWRYLL 580  
Db 1412 RKRKPYRPGIGGFMVQRKSHTRTKGPAQAQAEVLSDGQDPDEVIPADLPAGEAVEQSL- 1470  
QY 581 RRDDEPGPWTKEGDKRIKGLTQYPEGYLEALANRE----- 619  
Db 1471 -AEGDEKKKQRRGRKRSKLEG---FPAYLQEAFFGKELLDLSRKALFAVGVRPSFGL 1526

Qy 620 -----REKENSX-----EEEOGGFASPRGTGKWKKSAGGSPS 657  
Db 1527 GTPKAKGDDGSSERKELPTSQKDDGPDIADEESRGLEKADTPGEDGGVKASPVPSDFE 1586  
Qy 658 RAGSPRRTSKTKVEPYSLTAQSSLIREDKSNAKLWNEVLASLKDRLP-ASGSP 710  
Db 1587 KPGTPEGMLSSDLDRIS-TEELPKM--ESKDLQOLFVKDVLGSRERQHLGCGTP 1637

Search completed: November 1, 2004, 15:46:23  
Job time : 33 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:43:33 ; Search time 9 Seconds  
(without alignments)  
445.607 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 4272

Sequence: 1 MWIQRVTMDGRQTHVTDSLS.....VNQPLQTVLNQLFPFGYGNR 793

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 38294 seqs, 5057332 residues

Total number of hits satisfying chosen parameters: 38294

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_New\*

- 1: /cgn2\_6/prodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/prodata/2/paa/US05\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/prodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/prodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/prodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/prodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/prodata/2/paa/US11\_NEW\_COMB.pep.\*
- 8: /cgn2\_6/prodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	247	5.8	89	6	US-10-220-366A-21074
2	142	3.3	506	6	US-10-811-080-18
3	132.5	3.1	888	6	US-10-722-939-10
4	124.5	2.9	153	6	US-10-399-103A-482
5	124	2.9	76	1	PCT-US04-33137-10
6	124	2.9	76	6	US-10-835-096-13
7	124	2.9	76	6	US-10-067-832D-21
8	121	2.8	84	6	US-10-835-096-14
9	121	2.8	85	6	US-10-835-096-15
10	121	2.8	86	6	US-10-835-096-18
11	113	2.6	1356	1	PCT-US03-40884-20
12	108.5	2.5	405	8	US-60-613-154-50
13	108.5	2.5	1142	1	PCT-US04-14421-213
14	106	2.5	498	6	US-10-399-103A-616
15	106	2.5	3696	6	US-10-687-268-31
16	106	2.5	3705	6	US-10-687-268-30
17	103.5	2.4	353	6	US-10-955-952-296
18	103.5	2.4	353	6	US-10-157-779-296
19	103.5	2.4	353	6	US-10-964-241-296
20	98	2.3	36946	1	PCT-US04-14421-155
21	96	2.2	485	6	US-10-399-103A-628
22	95	2.2	366	6	US-10-510-386-42
23	95	2.2	366	1	PCT-US04-17765-62
24	94	2.2	1238	6	US-10-765-727-21
25	94	2.2	1238	6	US-10-846-989-55

ALIGNMENTS

RESULT 1  
US-10-220-366A-21074  
; Sequence 21074, Application US/10220366A  
; GENERAL INFORMATION:  
; APPLICANT: HYSEQ, INC  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-042  
; CURRENT APPLICATION NUMBER: US/10/220,366A  
; PRIOR FILING DATE: 2002-08-28  
; PRIOR APPLICATION NUMBER: 09/577,409  
; PRIOR FILING DATE: 2000-05-18  
; PRIOR APPLICATION NUMBER: 09/515,126  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 27802  
; SOFTWARE: Custom  
; SEQ ID NO 21074  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(89)  
; OTHER INFORMATION: Xaa = any amino acid or nothing  
US-10-220-366A-21074

Query Match 5.8%; Score 247; DB 6; Length 89;  
Best Local Similarity 55.0%; Pred. No. 8.8e-14;  
Matches 44; Conservative 15; Mismatches 21; Indels 0; Gaps 0;  
QY 714 ELKVEETFCQICQELVPRITTVQCNVCKDLRSFRAQVFCPCRYDLGRSYAMQ 773  
Db 10 FLTTMTCMLVCQELGYQVTTCHNVNCLQRFKAQVFSCHACHDLQNYIMI 69  
QY 774 VNQLPQTVLNQLFPFGYGNR 793  
Db 70 PNEIMQTLILDAPFGYSLGR 89

RESULT 2  
US-10-811-080-18  
; Sequence 18, Application US/10811080  
; GENERAL INFORMATION:  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Mehraban, Foad  
; TITLE OF INVENTION: NOVEL POLYPEPTIDES, THEIR NUCLEIC ACIDS, AND METHODS  
; FOR THEIR USE IN ANGIOGENESIS AND VASCULARIZATION



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 76
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
PCT-US04-33137-10

Query Match          2.9%; Score 124; DB 1; Length 76;
Best Local Similarity 37.8%; Pred. No. 0.00072;
Matches 28; Conservative 18; Mismatches 26; Indels 2; Gaps 2;

QY 1 MWIQVMTDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 1 MQIFVKLTGK-TITLE-VEPSTDIENVKAKIQDKGIPDPQQRLIFAGKQLEDGRTLSD 58
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY 61 YEVRNDTIQLLVR 74
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 59 YNIQESTLHLVLR 72
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 6
US-10-835-096-13
; Sequence 13, Application US/10835096
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Ubiquitin Ligase Assay
; FILE REFERENCE: 021044-007010US
; CURRENT APPLICATION NUMBER: US/10/835,096
; CURRENT FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/826,312
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 09/542,497
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human ubiquitin
US-10-835-096-13

Query Match          2.9%; Score 124; DB 6; Length 76;
Best Local Similarity 37.8%; Pred. No. 0.00072;
Matches 28; Conservative 18; Mismatches 26; Indels 2; Gaps 2;

QY 1 MWIQVMTDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 1 MQIFVKLTGK-TITLE-VEPSTDIENVKAKIQDKGIPDPQQRLIFAGKQLEDGRTLSD 58
   |||:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|

QY 61 YEVRNDTIQLLVR 74
   |||:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 59 YNIQESTLHLVLR 72
   |||:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|

RESULT 7
US-10-067-832D-21
; Sequence 21, Application US/10067832D
; GENERAL INFORMATION:
; APPLICANT: ZIMMET, PAUL ZEV
; APPLICANT: COLLIER, GREGORY
; TITLE OF INVENTION: A NOVEL GENE AND USES THEREFOR
; FILE REFERENCE: 229752000701
; CURRENT APPLICATION NUMBER: US/10/067,832D
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 09/331,930
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: PCT/AU98/00902
```

```
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: AU PP 0117
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: AU PP 0323
; PRIOR FILING DATE: 1997-11-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 21
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-832D-21

Query Match          2.9%; Score 124; DB 6; Length 76;
Best Local Similarity 37.8%; Pred. No. 0.00072;
Matches 28; Conservative 18; Mismatches 26; Indels 2; Gaps 2;

QY 1 MWIQVMTDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60
   |||:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 1 MQIFVKLTGK-TITLE-VEPSTDIENVKAKIQDKGIPDPQQRLIFAGKQLEDGRTLSD 58
   |||:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|

QY 61 YEVRNDTIQLLVR 74
   |||:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 59 YNIQESTLHLVLR 72
   |||:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|

RESULT 8
US-10-835-096-14
; Sequence 14, Application US/10835096
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Ubiquitin Ligase Assay
; FILE REFERENCE: 021044-007010US
; CURRENT APPLICATION NUMBER: US/10/835,096
; CURRENT FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/826,312
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 09/542,497
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: FLAG-ubiquitin
US-10-835-096-14

Query Match          2.8%; Score 121; DB 6; Length 84;
Best Local Similarity 37.5%; Pred. No. 0.0015;
Matches 27; Conservative 18; Mismatches 25; Indels 2; Gaps 2;

QY 3 IQVRTMDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFDYE 62
   |||:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 11 IFVKLTGK-TITLE-VEPSTDIENVKAKIQDKGIPDPQQRLIFAGKQLEDGRTLSDYN 68
   |||:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|

QY 63 VRLNDTIQLLVR 74
   |||:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 69 IQKESTLHLVLR 80
   |||:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|

RESULT 9
US-10-835-096-15
; Sequence 15, Application US/10835096
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
```

```

; APPLICANT: Pray, Todd R.
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Ubiquitin Ligase Assay
; FILE REFERENCE: 021044-007010US
; CURRENT APPLICATION NUMBER: US/10/835,096
; CURRENT FILING DATE: 2004-04-28
; PRIOR FILING DATE: 2001-04-03
; PRIOR FILING DATE: 2001-04-03
; PRIOR FILING DATE: 2001-04-03
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial
; OTHER INFORMATION: Sequence: FLAG-Cys-ubiquitin
US-10-835-096-15

Query Match 2.8%; Score 121; DB 6; Length 85;
Best Local Similarity 37.5%; Pred. No. 0.0015;
Matches 27; Conservative 18; Mismatches 25; Indels 2; Gaps 2;

QY 3 IQVMTDGRQTHVTDSLSRLTKVEELRRKIQELFHFVEGLQRLFYRGKQMEDGHTLFDYE 62
Db 12 IFVKTGK-TITLE-VEPSTIENVKAKIQKEGIPDPQORLIFAGKQLEDGRTLSDYN 69

QY 63 VRLNDTIQLLVR 74
Db 70 IQKESTLHLVLR 81

RESULT 10
US-10-835-096-18
; Sequence 18, Application US/10835096
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Ubiquitin Ligase Assay
; FILE REFERENCE: 021044-007010US
; CURRENT APPLICATION NUMBER: US/10/835,096
; CURRENT FILING DATE: 2004-04-28
; PRIOR FILING DATE: 2001-04-03
; PRIOR FILING DATE: 2001-04-03
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial
; OTHER INFORMATION: Sequence: FLAG-Ala-Cys-ubiquitin
US-10-835-096-18

Query Match 2.8%; Score 121; DB 6; Length 86;
Best Local Similarity 37.5%; Pred. No. 0.0015;
Matches 27; Conservative 18; Mismatches 25; Indels 2; Gaps 2;

QY 3 IQVMTDGRQTHVTDSLSRLTKVEELRRKIQELFHFVEGLQRLFYRGKQMEDGHTLFDYE 62
Db 13 IFVKTGK-TITLE-VEPSTIENVKAKIQKEGIPDPQORLIFAGKQLEDGRTLSDYN 70

QY 63 VRLNDTIQLLVR 74
Db 71 IQKESTLHLVLR 82
```

RESULT 11

```

PCT-US03-40884-20
; Sequence 20, Application PC/TUS0340884
; GENERAL INFORMATION:
; APPLICANT: Sequenom, Inc.
; APPLICANT: Langdown, Maria L.
; APPLICANT: Nelson, Matthew Roberts
; APPLICANT: Reneland, Rikard Henry
; APPLICANT: Kammerer, Stefan M.
; APPLICANT: Braun, Andreas
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF TYPE II
; TITLE OF INVENTION: DIABETES AND TREATMENTS THEREOF
; FILE REFERENCE: 524592007640
; CURRENT APPLICATION NUMBER: PCT/US03/40884
; CURRENT FILING DATE: 2003-12-22
; PRIOR APPLICATION NUMBER: 60/435,431
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/499,143
; PRIOR FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: 60/498,100
; PRIOR FILING DATE: 2003-08-26
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1356
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-40884-20

Query Match 2.6%; Score 113; DB 1; Length 1356;
Best Local Similarity 20.1%; Pred. No. 0.37;
Matches 164; Conservative 109; Mismatches 298; Indels 246; Gaps 38;

QY 3 IQVMTD--GRQTHVTDSLSRLTKVEE-----LRKKIQELFH-----VEPGLQRLFYRKQ 51
Db 512 IEVNGVDLVGKSQEVEWSLLRSTKMEGTVSLLVFRQEDAFHPRELNAEPQMQIPKTKA 571

QY 52 MEDGHTL-----FDYEVLNDTIQLLVQSLVPLHSTKERDSELSDTSGCCLGQSE 103
Db 572 EDEDIVLTPDGTREFLTFEVLNDS-----GSAGLVSVKGNRSHKHAHLGIFV---- 621

QY 104 SDKSSTHGEAAAEITDSRPADMDMDETLGLYKNEYVDARDTNMGANFAEQVVRTRKA 163
Db 622 ---KSIINGAASKD-----GRLRVND-----QLIAV----- 645

QY 164 PSRDEPCSTSRPALLEDVIYHVYDDYPENGVOVMNSRDVRRARTIKWQDLEVGQV 223
Db 646 -NGESLLGKTNQDAME--TLRRSMTEGNGRGMQL-----IVARRISKCNELK----- 691

QY 224 MLNYPNDNPKRGFWYDAEISRKRETRTARELYANVLGDDSLNDCRIIFVDEVEKIERP 283
Db 692 ----SPGSPGPPELPIETALD--DRERRISHSLYS-----GIEGLDE----- 727

QY 284 GEGSPMVDNPMRRKSGPCKH--CKDDVNRLCRVCACHLCGGRRQDPDKQLMCDCECDMAFI 342
Db 728 ----SPSRNAALSRIMGESCKQLSFTVM-----PQDDTVIIEDD----RL 766

QY 343 YCLDPPL---SSVPSEDE-----WYCEPCNDAS-----EVLAGERL---RE 379
Db 767 PVLPPHLSQSSSSSHDDVGFVTADAGTWAKAAISDSADCSLSPDVPDLAPQRFQFGFQ 826

QY 380 SKKNAKMASATSSSQORDWCK-----GMACVGRITKECTI-----VPSNHYGPPIGI-- 424
Db 827 SMSEKETQFSDASQDLDFVKTTRKSKMDLGIADETKLNITVDQKAGSPRDVPSLGLKK 886

QY 425 --PVGTWRRFRVQVSESG---VHRPHVAGIHGRSNDGYSYLVLAGYE-----DDVDHGNP 475
Db 887 SSSLESLSQTAVAEVLINGDI PFHRPRPRIIRGCGNESFRAAIDKSYDKPAVDDDDDEGME 946

QY 476 FTYTSGGRDLGSKNKRTAEQSCDQKLTNTNRLALNCFAPINDQEGAEAKDWRSGKPVV 535
```

```
Db 947 TLEBDTESSRSRESYSTAS-DQPSHSLERQMNGN-----QKGDKTDKDK----- 994
QY 536 VRNVKGGKNSKXAPAEGRYDGIYKVWPEKKGSGFL-----VMRYLLRRDDDFGPW 590
Db 995 -----TGKEKKKDRDKDKM-----KAKGMLKGLGDMFRFGKHKDKDK--I 1035
QY 591 TKEGKDKIKKL-----GLTWOYPEGYLEALANREKENSKEEBEQO----- 633
Db 1036 EKTGKIKIQSFTEEBIRIMKQBERIQAKTRFRERQARERYABIQDFHRTFGCDDE 1095
QY 634 -----EGGFA-----SPRTGK-----GKWKRSAGGSPRAGSPRRT-SKTKVE 672
Db 1096 LMVGGVSYEGSMALNAPQSPREGHMDALYAQVKPRNSKPPVDSNRSFNSHNRIO 1155
QY 673 PYSITAQOSSLIRDKSNAKLWNEVLASLKDORPASGS 709
Db 1156 RLROFQOAKODEVEDRRRTYSFEPWPNARPATQS 1192

RESULT 12
US-60-613-154-50
; Sequence 50, Application US/60613154
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science
; TITLE OF INVENTION: PLANT CELLS AND PLANTS WITH INCREASED TOLERANCE TO ENVIRONMENTAL
; FILE REFERENCE: STRESS
; CURRENT APPLICATION NUMBER: US/60/613,154
; CURRENT FILING DATE: 2004-09-24
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-60-613-154-50

Query Match 2.5%; Score 108.5; DB 8; Length 405;
Best Local Similarity 28.4%; Pred. No. 0.15;
Matches 42; Conservative 27; Mismatches 62; Indels 17; Gaps 8;

QY 651 SAGGSP-SRAGSPRRTSKTKVPPYSITAQOSSLIRDKSNAKLWNEVLASL-KDRPASG 708
Db 27 TAGGCFESKAARDPDDASARKQGETTAGGCFPSKARPENSGKCFEQEGLNKNOSTDSA 86
QY 709 S-PFQLFLSKVETFO-----CICQELVPRITTV-COHNVKOCLDRSFRAQVFSCEA 761
Db 87 TVPAKCPFGYDSQTFKLGFPFSCMLQALLYESRRCVPCPTHVFCVKCLTR-FK-----DCPL 141
QY 762 CRYDLGRSYAMOVNOLQTVLNLFPGY 789
Db 142 CGADI-----ESIEVDENLQXWVDFIEGH 166

RESULT 13
PCT-US04-14421-213
; Sequence 213, Application PC/TUS0414421
; GENERAL INFORMATION:
; APPLICANT: SUGEN, INC.
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: MANNING, GERARD
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: GRIGORIEV, IGOR
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-1454
; CURRENT APPLICATION NUMBER: PCT/US04/14421
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/469,014
; NUMBER OF SEQ ID NOS: 239
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 213
; LENGTH: 1142
```

```
; TYPE: PRT
; ORGANISM: Mus musculus
PCT-US04-14421-213

Query Match 2.5%; Score 108.5; DB 1; Length 1142;
Best Local Similarity 20.6%; Pred. No. 0.68;
Matches 74; Conservative 45; Mismatches 145; Indels 95; Gaps 15;

QY 51 QMEDGHTLFQVEYRLNDTIQLLVRSQSLVLPKSTKERD-----SELSDTDSGCLGQSES- 104
Db 702 QLEDAGS-----SSLDNLLSRYSGLPQPTSTWMPSPGSPALSFGSG--LSNSHTP 754
QY 105 -----DKSSTHGEAAETDSRPADENMDETELGLYKVNXYVDARTNMGAWFAEQVVRVT 160
Db 755 VRPPTSTSTGSRGSGSGSRTAEKSAHS-----FKSDQVKVK 791
QY 161 RKAPSRDEPCSSSTRPALEEDVIYHVKYDDYPENGVVOMNSRDVVRARARTIIKWDLEVG 220
Db 792 QEPGTEEEICFS-----GAVKQEKTEGRRSA----- 819
QY 221 QVVMNPNPNPKERGFWDYDAEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKI 280
Db 820 --CMLS-SPESLTPPLSTNLHLESLDTLTGLNHHVKTPTDIS-ESCKQSGLSNLV-- 873
QY 281 ERPEGSPVVDNPMRRKS--GPSCKHKCKDDVNR-LCRVCACHLCGGRQDPDKQCMDECD 337
Db 874 ----NGKSPIRNLHRSARIGDGNKDDDPNEDWCAVC-----QNGGDLCCCKCP 921
QY 338 MAFHIYCLDPPLSSVSEDEWYCEPCRN-DASVVLAGERLRESKNAKVASATSSOR 395
Db 922 KVPHLTCHVPTLLSFPSGD-WICTFCRDIGKPEVEYDCDNMOMHSGKKGKTAQGLSPVDQR 979

RESULT 14
US-10-399-103A-616
; Sequence 616, Application US/10399103A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-115/785
; CURRENT APPLICATION NUMBER: US/10/399,103A
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: PCT/US01/27760
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 09/687,527
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 892
; SOFTWARE: Custom
; SEQ ID NO 616
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-399-103A-616

Query Match 2.5%; Score 106; DB 6; Length 498;
Best Local Similarity 24.5%; Pred. No. 0.32;
Matches 27; Conservative 19; Mismatches 52; Indels 12; Gaps 5;

QY 282 RPGEGPSVVDNPMRRKSG-PSCKHKCKDDVNR-----CRVC-ACHLCGGRQDPDKQLM 332
Db 391 KKGKAESLIHSCQENSGHPSCLDMTLMELVSMIKTYPWQCMCKCTCIIGQPHHEEMMF 450
QY 333 CDECDMAFHIYCLDPPLSSVSEDEWYCEPCRNDASEVVVLAGERLRESKK 382
Db 451 CDMCDRGYHTFCVG--LGAIPS-GRWICCCORAPPTPRKVGRRGKNSKE 497

RESULT 15
US-10-687-268-31
; Sequence 31, Application US/10687268
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Lee, Judithann M.
```

```

; APPLICANT: Smith, Randall F.
; APPLICANT: White, John R.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50029-1
; CURRENT APPLICATION NUMBER: US/10/687,268
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 10/312,088
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 3696
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-687-268-31

Query Match          2.5%; Score 106; DB 6; Length 3696;
Best Local Similarity 18.6%; Pred. No. 5.9;
Matches 121; Conservative 45; Mismatches 180; Indels 304; Gaps 30;

QY      238 WYDAETSRKRETTARELYANVGLDSDLNDCRIIFVDEPKIEREGSGSPVMDNPMRKK 297
Db      369 YYDEPVDRRRASQSLDGTQY-----GGGVCIDCQ-----HHT 400

QY      298 SGPSCXHKCD-----DVNRLRCVCACH---LGGRODDPKQLMC-----DECDM 338
Db      401 TGVNCRCLPGFYRSPNHPLDSPHVCRNCESDFTDGTCEDLTGRCYCRFNFSGERCDV 460

QY      339 AFHYICLDPLSSVPSEDEWYPCERNDASEVVL-AGERLRSEKKNAMKASATSSSQDWM 397
Db      461 CAEGFTGFPSCYPTPSSS-----NDTREQVLPAGQIV-----NCDCSAAGT----- 501

QY      398 GKGMAC-----VGRT-----KECTIVPSNHYGP-----IPGI----- 424
Db      502 -QGNACKRKDRVGRCLCKPNFQGHCELCAPGFTYGGQPCQCQSSPVADDRCDPDGTQC 560

QY      425 -----PVGTM-----WRFVQVSESGVH-----R 443
Db      561 RCRVGFEGATCDRCAPGFHPLCQLCGCSPAGTLPFGCEAGRCCLCOPEFAGPHCDRCR 620

QY      444 PHVAGIHGRNDGSYSLVLAGGYEDDVHGNFFYTGSGG--RDLSGNKRTAEQSCDQKL 501
Db      621 P--GVHGFPNCOACTCDPRGALDQ-----LCGAGGLCRCPGVTGTACQECSPGF 668

QY      502 TINTNRALALNCFAPINDOEGA--EAKOWRSKG-----PVRVVRNVKGGKNSKYAPAE 551
Db      669 HGPSPSCVPCHCSA-----EGSLHAACDRPSQCSCRPVTVGLRCDTCVPGAYNFFYCRAG 723

QY      552 GNRVDGYIKVVKYKWE-----KKGSGFLVMRYLLRRDDDEPGPWTKEGKDRIKK 600
Db      724 SCHPAGLAPVDPALPEAQVPCWCMRAHVEGPS-----CDRCKPGFW----- 763

QY      601 LGLTMQVPEGYLEALANRERKENSKRREEEQEGGFASPTGKKGKWKRKSKAGGFSRAG 660
Db      764 -GLSPSPNEPGCTRCSIDLRTGL-----GGVAECQPGTG----- 795

QY      661 SPRTSKTKVEPYSLTAQSSLTIREDKSNAKLWNEVLASLAKDRPASGPFQLFLSKVEE 720
Db      796 ----- 795

QY      721 TFQCTCCOELVFRPITTVQHNVCVKCDLRSF---RAQVFSQPCARYDLG 767
Db      796 --QCFC-----XP--HYC--GOACASCKDGFGLDQADYFGCRSCRCIDIG 834

```

Search completed: November 1, 2004, 15:50:07  
Job time : 10 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:45:23 ; Search time 83 Seconds

(without alignments)

3097.628 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 4272

Sequence: 1 MWIQRVTMDGRQTHVTDSL.....VNQPLQTLNQLPPGYGNR 793

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/CTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09C\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4263	99.8	793	14	US-10-123-568-2
2	4263	99.8	793	15	US-10-188-832-27
3	4191	98.1	780	14	US-10-295-027-156
4	2293.5	53.7	802	16	US-10-476-924-7
5	1872	43.8	645	14	US-10-126-103-113
6	1872	43.8	645	15	US-10-431-096-113
7	823.5	19.3	198	9	US-09-867-550-766
8	736	17.2	136	14	US-10-123-568-3
9	672	15.7	133	9	US-09-764-864-1301
10	626.5	14.7	694	15	US-10-425-114-39797
11	626.5	14.7	709	15	US-10-424-599-174378
12	588.5	13.8	617	14	US-10-225-066A-522
13	588.5	13.8	617	15	US-10-374-780A-2314

14	498	11.7	750	16	US-10-437-963-156872	Sequence 156872,
15	478.5	11.2	774	16	US-10-437-963-180766	Sequence 180766,
16	420.5	9.8	178	9	US-09-764-864-1303	Sequence 1303, Ap
17	357.5	8.4	110	9	US-09-764-864-848	Sequence 848, App
18	331	7.7	438	15	US-10-424-599-263042	Sequence 263042,
19	279	6.5	684	16	US-10-437-963-132977	Sequence 132977,
20	270	6.3	810	15	US-10-425-114-54293	Sequence 54293, A
21	266.5	6.2	273	15	US-10-425-114-41693	Sequence 41693, A
22	236	5.5	707	15	US-10-424-599-277747	Sequence 277747,
23	232.5	5.4	557	15	US-10-424-599-167978	Sequence 167978,
24	230.5	5.4	682	16	US-10-437-963-109882	Sequence 109882,
25	229.5	5.4	600	15	US-10-425-114-57455	Sequence 57455, A
26	229.5	5.4	856	14	US-10-310-154-600	Sequence 600, App
27	227	5.3	297	15	US-10-424-599-171858	Sequence 171858,
28	223	5.2	678	16	US-10-437-963-145990	Sequence 145990,
29	221	5.2	768	16	US-10-437-963-133503	Sequence 133503,
30	219.5	5.1	396	15	US-10-425-114-70940	Sequence 70940, A
31	219	5.1	250	16	US-10-767-701-33924	Sequence 33924, A
32	216.5	5.1	673	16	US-10-437-963-159376	Sequence 159376,
33	215.5	5.0	730	15	US-10-425-114-72673	Sequence 72673, A
34	214	5.0	1292	16	US-10-437-963-136868	Sequence 136868,
35	213.5	5.0	812	16	US-10-437-963-136868	Sequence 136868,
36	212.5	5.0	329	16	US-10-767-701-41660	Sequence 41660, A
37	209.5	4.9	400	16	US-10-437-963-159375	Sequence 159375,
38	204.5	4.8	999	16	US-10-437-963-164255	Sequence 164255,
39	195	4.6	1358	16	US-10-437-963-112581	Sequence 112581,
40	193.5	4.5	299	15	US-10-424-599-148375	Sequence 148375,
41	193.5	4.5	1398	14	US-10-094-466-30	Sequence 30, Appl
42	193.5	4.5	1400	9	US-09-764-176-7	Sequence 7, Appl
43	191.5	4.5	178	16	US-10-437-963-180768	Sequence 180768,
44	178.5	4.2	290	15	US-10-425-114-36682	Sequence 36682, A
45	178.5	4.2	387	15	US-10-424-599-242736	Sequence 242736,

#### ALIGNMENTS

RESULT 1  
US-10-123-568-2  
; Sequence 2, Application US/10123568  
; Publication No. US20030194713A1  
; GENERAL INFORMATION:  
; APPLICANT: Hitoshi, Yasumichi  
; APPLICANT: Jenkins, Yonchu  
; APPLICANT: Rigel Pharmaceuticals, Inc.  
; TITLE OF INVENTION: NP95: Methods of Assaying for Cell Cycle Modulators  
; FILE REFERENCE: 021044-003400US  
; CURRENT APPLICATION NUMBER: US/10/123,568  
; CURRENT FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 793  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human NP95 nuclear zinc finger protein  
US-10-123-568-2

Query Match	99.8%	Score 4263;	DB 14;	Length 793;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 791;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MWIVRTMDGRQTHVTDSLRLTKVEELRRKIQELFVEPGLQRLFYRGKQMGHTILFD	60	
Db	1	MWIVRTMDGRQTHVTDSLRLTKVEELRRKIQELFVEPGLQRLFYRGKQMGHTILFD	60	
Qy	61	YEVLNDTIQLLRQSLVLPSTKERSLSDTSGCGLQSESDKSTHGEAAETDSR	120	
Db	61	YEVLNDTIQLLRQSLVLPSTKERSLSDTSGCGLQSESDKSTHGEAAETDSR	120	
Qy	121	PAEDMDDETLGLYKNEVVDARDTNMGAFQVVRTRKAFSRDEPCSSSRPALEE	180	

```

121 PADEDMWDETELGLYKVNVEYDARDTNMGAWFAQVVRVTRKAPSRDEPCSSSTRPALEE 180
181 DVYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPNPNKRGFWYD 240
181 DVYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPNPNKRGFWYD 240
241 AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGESEPMVDNPMRRKSGP 300
241 AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGESEPMVDNPMRRKSGP 300
301 SKCHKDDVNRLCRVCAHLGGRQDDPKQIMCDECDMAFHYICLDPLSPVSEDEWYC 360
301 SKCHKDDVNRLCRVCAHLGGRQDDPKQIMCDECDMAFHYICLDPLSPVSEDEWYC 360
361 PECDNDASEVVLAGERLRESKKNAMASATSSORDMGKMACVGRTECTIVPSNHYGP 420
361 PECDNDASEVVLAGERLRESKKNAMASATSSORDMGKMACVGRTECTIVPSNHYGP 420
421 IPGIPVGTMRFRVQVSESGVHRPHVAGIHRGSDGYSILVLAGGYEDDVGHNFFTYTG 480
421 IPGIPVGTMRFRVQVSESGVHRPHVAGIHRGSDGYSILVLAGGYEDDVGHNFFTYTG 480
481 SGGDLSCNKRRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540
481 SGGDLSCNKRRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540
541 GGNKSKYAPAEGRNYDGIYKVKYWPKEKSGFLVWRYLLRRDDDEPGPMTKEGDRICK 600
541 GGNKSKYAPAEGRNYDGIYKVKYWPKEKSGFLVWRYLLRRDDDEPGPMTKEGDRICK 600
601 LGLTMQYPEGYLEALANRERKENSKEEBEQEGGFASPRTGKWKRSAGGSPSRAG 660
601 LGLTMQYPEGYLEALANRERKENSKEEBEQEGGFASPRTGKWKRSAGGSPSRAG 660
721 TFOCICCOBELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCACRYDLGRSYAMQVNOPIQT 780
721 TFOCICCOBELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCACRYDLGRSYAMQVNOPIQT 780
781 VLNLFPFGYNGNR 793
781 VLNLFPFGYNGNR 793

```

RESULT 2

```

US-10-188-832-27
; Sequence 27, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natacha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 27
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-27

```

Query Match 99.8%; Score 4263; DB 15; Length 793;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 791; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MWIQVTRTMDGROTHTVDSLSRLTKVBELRRKIIQELFHVPEGLQRLFRGKQMEDGHTLFD 60
DB 1 MWIQVTRTMDGROTHTVDSLSRLTKVBELRRKIIQELFHVPEGLQRLFRGKQMEDGHTLFD 60
QY 61 YEVRLNDTIQLLRQSLVLPKSTKERSELSDDTSGCCLGQSESDKSTHGEAAAEADSR 120
DB 61 YEVRLNDTIQLLRQSLVLPKSTKERSELSDDTSGCCLGQSESDKSTHGEAAAEADSR 120
QY 121 PADEDMWDETELGLYKVNVEYDARDTNMGAWFAQVVRVTRKAPSRDEPCSSSTRPALEE 180
DB 121 PADEDMWDETELGLYKVNVEYDARDTNMGAWFAQVVRVTRKAPSRDEPCSSSTRPALEE 180
QY 181 DVYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPNPNKRGFWYD 240
DB 181 DVYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPNPNKRGFWYD 240
QY 241 AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGESEPMVDNPMRRKSGP 300
DB 241 AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGESEPMVDNPMRRKSGP 300
QY 301 SKCHKDDVNRLCRVCAHLGGRQDDPKQIMCDECDMAFHYICLDPLSPVSEDEWYC 360
DB 301 SKCHKDDVNRLCRVCAHLGGRQDDPKQIMCDECDMAFHYICLDPLSPVSEDEWYC 360
QY 361 PECDNDASEVVLAGERLRESKKNAMASATSSORDMGKMACVGRTECTIVPSNHYGP 420
DB 361 PECDNDASEVVLAGERLRESKKNAMASATSSORDMGKMACVGRTECTIVPSNHYGP 420
QY 421 IPGIPVGTMRFRVQVSESGVHRPHVAGIHRGSDGYSILVLAGGYEDDVGHNFFTYTG 480
DB 421 IPGIPVGTMRFRVQVSESGVHRPHVAGIHRGSDGYSILVLAGGYEDDVGHNFFTYTG 480
QY 481 SGGDLSCNKRRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540
DB 481 SGGDLSCNKRRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540
QY 541 GGNKSKYAPAEGRNYDGIYKVKYWPKEKSGFLVWRYLLRRDDDEPGPMTKEGDRICK 600
DB 541 GGNKSKYAPAEGRNYDGIYKVKYWPKEKSGFLVWRYLLRRDDDEPGPMTKEGDRICK 600
QY 601 LGLTMQYPEGYLEALANRERKENSKEEBEQEGGFASPRTGKWKRSAGGSPSRAG 660
DB 601 LGLTMQYPEGYLEALANRERKENSKEEBEQEGGFASPRTGKWKRSAGGSPSRAG 660
QY 661 SPRTSKTKVPEYSLTAQOSSLIREDKSNKLNWNEVLASLKDPAASGSPFQLFLSKVEE 720
DB 661 SPRTSKTKVPEYSLTAQOSSLIREDKSNKLNWNEVLASLKDPAASGSPFQLFLSKVEE 720
QY 721 TFOCICCOBELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCACRYDLGRSYAMQVNOPIQT 780
DB 721 TFOCICCOBELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCACRYDLGRSYAMQVNOPIQT 780
QY 781 VLNLFPFGYNGNR 793
DB 781 VLNLFPFGYNGNR 793

```

RESULT 3

```

US-10-295-027-156
; Sequence 156, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel

```

APPLICANT: Aziz, Natasha  
APPLICANT: Ginsberg, Wendy M.  
APPLICANT: Gish, Kurt C.  
APPLICANT: Glynn, Richard  
APPLICANT: Hevezil, Peter A.  
APPLICANT: Mack, David H.  
APPLICANT: Murray, Richard  
APPLICANT: Watson, Susan R.  
APPLICANT: Eos Biotechnology, Inc.  
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and Methods of Screening for Modulators of Cancer  
FILE REFERENCE: 018501-012500US  
CURRENT APPLICATION NUMBER: US/10/295,027  
CURRENT FILING DATE: 2002-11-13  
PRIOR APPLICATION NUMBER: US 09/663,733  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: US 60/350,666  
PRIOR FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 60/335,394  
PRIOR FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: US 60/332,464  
PRIOR FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: US 60/334,393  
PRIOR FILING DATE: 2001-11-29  
PRIOR APPLICATION NUMBER: US 60/340,376  
PRIOR FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: US 60/347,211  
PRIOR FILING DATE: 2002-01-08  
PRIOR APPLICATION NUMBER: US 60/347,349  
PRIOR FILING DATE: 2002-01-10  
PRIOR APPLICATION NUMBER: US 60/355,250  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: US 60/356,714  
PRIOR FILING DATE: 2002-02-13  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1386  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 156  
LENGTH: 780  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-295-027-156

Query Match 98.1%; Score 4191; DB 14; Length 780;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 778; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MWIQRWTDGRQTHVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60  
DB 1 MWIQRWTDGRQTHVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60  
QY 61 YEVRINDTIQLLVROSLVLPSTKRDSELSLSDTSGCLGQSESDKSTHGAAETDSR 120  
DB 61 YEVRINDTIQLLVROSLVLPSTKRDSELSLSDTSGCLGQSESDKSTHGAAETDSR 120  
QY 121 PADEMDMETELGLYKNEYVDARTNNGANFEAOVVRVTRKAPSRDEPCSTSRPALEE 180  
DB 121 PADEMDMETELGLYKNEYVDARTNNGANFEAOVVRVTRKAPSRDEPCSTSRPALEE 180  
QY 181 DVIYHVYKDDYPENGVOVMSRDRVARARTTIKQDLEVGQVVMNPNPNKRGFWYD 240  
DB 181 DVIYHVYKDDYPENGVOVMSRDRVARARTTIKQDLEVGQVVMNPNPNKRGFWYD 240  
QY 241 AEISRKRETRARELYANVLGDSLNDCLRIIFVDEVFKEIRPGEPSWMDPNMRKSGP 300  
DB 241 AEISRKRETRARELYANVLGDSLNDCLRIIFVDEVFKEIRPGEPSWMDPNMRKSGP 300  
QY 301 SCRKCDVNLRCVACHLCGGQDPDPKQMLCDECDMAEHYICLDPLSSVPSDEWYC 360  
DB 301 SCRKCDVNLRCVACHLCGGQDPDPKQMLCDECDMAEHYICLDPLSSVPSDEWYC 360  
QY 361 PECNDASEVVLAGERLRESKKAAMASATSSSQRDWKGMACVGRTECTIVPSNHVGP 420

DB 361 PECNDASEVVLAGERLRESKKAAMASATSSSQRDWKGMACVGRTECTIVPSNHVGP 420  
QY 421 IPGIPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSILVLAGYEDDVHGNFFYVTG 480  
DB 421 IPGIPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSILVLAGYEDDVHGNFFYVTG 480  
QY 481 SGGRLSGNKRRTABQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540  
DB 481 SGGRLSGNKRRTABQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540  
QY 541 GGNKSKYAPAEGRNYDGIYKVYKWPBKSGFLVWRYLLRRDDDEPGPWTKEGDKRIK 600  
DB 541 GGNKSKYAPAEGRNYDGIYKVYKWPBKSGFLVWRYLLRRDDDEPGPWTKEGDKRIK 600  
QY 601 LGLTMQYPEGYLEALANREREREKREKREKREKREKREKREKREKREKREKREKREK 660  
DB 601 LGLTMQYPEGYLEALANREREREKREKREKREKREKREKREKREKREKREKREKREK 660  
QY 661 SPRTSKTKVPEYSLTAQSSSLIREDKSNALWNEVLASLKDPSAGSGSPFOLFLSKVEE 720  
DB 661 SPRTSKTKVPEYSLTAQSSSLIREDKSNALWNEVLASLKDPSAGSGSPFOLFLSKVEE 720  
QY 721 TFQCICQBELVRPITTVCOHNVCKDCLDRSFRAQVFSQPCRYDLGRSYAMQVNPLOT 780  
DB 721 TFQCICQBELVRPITTVCOHNVCKDCLDRSFRAQVFSQPCRYDLGRSYAMQVNPLOT 780

RESULT 4  
US-10-476-924-7  
Sequence 7, Application US/10476924  
Publication No. US20040152093A1  
GENERAL INFORMATION:  
APPLICANT: YUE, Henry; DING, Li;  
APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;  
APPLICANT: YUE, Huibin; HAFALIA, April J.A.;  
APPLICANT: LEE, Ernestine A.; ISON, Craig H.;  
APPLICANT: BECHA, Shanya D.; GURURAJAN, Rajagopal;  
APPLICANT: EMERLING, Brooke M.; GRIFFIN, Jennifer A.;  
APPLICANT: TANG, Y. Tom; LU, Dyung Aina M.;  
APPLICANT: YAO, Monique G.; CHAWLA, Narinder K.;  
APPLICANT: RAMKUMAR, Jayalaxmi; GANDHI, Ameena R.;  
APPLICANT: YANG, Junming; ELLIOTT, Vicki S.;  
APPLICANT: LU, Yan; THANGAVELU, Kavitha;  
APPLICANT: HE, Ann; AZIMZAI, Valda;  
APPLICANT: RAUMANN, Brigitte E.; SWARNAKAR, Anita;  
APPLICANT: BURFORD, Neil  
TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS  
FILE REFERENCE: PF-0960 USN  
CURRENT APPLICATION NUMBER: US/10/476,924  
CURRENT FILING DATE: 2003-11-04  
PRIOR APPLICATION NUMBER: PCT/US02/14276  
PRIOR FILING DATE: 2002-05-02  
PRIOR APPLICATION NUMBER: US 60/288,598  
PRIOR FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: US 60/291,776  
PRIOR FILING DATE: 2001-05-17  
PRIOR APPLICATION NUMBER: US 60/292,172  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/293,564  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PERL Program  
SEQ ID NO 7  
LENGTH: 802  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No: 4706628CD1  
US-10-476-924-7

Query Match 53.7%; Score 2293.5; DB 16; Length 802;

Tue Nov 2 11:31:39 2004

Best Local Similarity 53.3%; Pred. No. 4.6e-178;  
Matches 444; Conservative 126; Mismatches 192; Indels 71; Gaps 14;  
QY 1 MWIQTMTGROQTHVDSLSRLTKVELRRKIQELHVEPGLQRLFRYRGKQMEDGHTLFD 60  
Db 1 MWIQTMTGROQTHVDSLSRLTKVELRRKIQELHVEPGLQRLFRYRGKQMEDGHTLFD 60  
QY 61 YEVRLNDTIQLLRQSL-VLPHSTKRDSELSLSDTSGCCLGQSDSKSSTHGEAAETDS 119  
Db 61 YDVLNDIIQLLRPDHLPFGTSTQIEAK-----PCSNSPPKVKAPRVGSPNQPSTS 114  
QY 120 RPADEDMWDETEGLYKVNVEVDARDTNMGAWFEAQQVVRVTR-----KAPSRD---- 167  
Db 115 ARA---RLIDPGFGIYKVNELVDARDVGLGAWFEAHIHVSSTRASDQSGRGKTPLNKSSC 171  
QY 168 -----EPCSSSRP---ALBEDVIYHVKYDDYPNGVVMQNSRDVRA 206  
Db 172 KRTNGNIKHSKENTNKLDSPVSTNSDCVAADSDVIYHIQYDEYPSGTLNEMVVKDLRP 231  
QY 207 RARTIIKQWQLEVGQVVMNPNPNKXERGFWDYDAEISR-KRETRTARELYANVVLG--D 263  
Db 232 RARTILKNELNVGDVVMVNVNVPSPGQGFWDYDAEITTLKTISRTKKELRVKIFLGSE 291  
QY 264 DSLNDCRIIFVDEYFKIERPGESEPM--VDNPMRKSPGSKCHKDDVNLRCVCAHLG 321  
Db 292 GTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLERNDPECDLGGDPEKKHSCSRVC 350  
QY 322 GGRQDPKQMLCDMCDMAFIYCLDPLSSVPSDEWYCPCEKNDASVVLGELRLRESK 381  
Db 351 GKGHEPNMQLCDECNVAIHYICLNPDLKVPBEEYWCPSCKTDSSEVVKAGERLKWSK 410  
QY 382 KNAKMASATSSQORDWKGMACVGRTECTIVPSNHYGPIPGIPVGTWRRFRVQVSEGV 441  
Db 411 KKAQPSASTESRRDWRGMACVGRTECTIVPSNHYGPIPGIPVGTWRRFRVQVSEAGV 470  
QY 442 HRPHVAGHGRSNDGYSILVLAGYEDVDHGNFTYTGSGGRDLSGNKRKTAEOQCDOKL 501  
Db 471 HRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSAQTL 530  
QY 502 TMTNPAALNCFAPINDQEGAEKDWRSKGPVVRVNVNKGKNSKYAPAEAGNRYDGIYKV 561  
Db 531 TMTNPAALNCDAPLDDKIGAESRNWRACKPVRVIRSFKGRKISKYAPAEAGNRYDGIYKV 590  
QY 562 VKYWEKGS-KFLVWRYLLRDRDDDEPGWTKGKDRIKKGL 603  
Db 591 VKYWEISSHGHFLVWRYLLRDRDDVEPAPWTSEGIERSRRLCL 633

RESULT 5  
US-10-126-103-113  
; Sequence 113, Application US/10126103  
; Publication No. US20030224486A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB PATHWAY  
; FILE REFERENCE: D0108.np  
; CURRENT APPLICATION NUMBER: US/10/126,103  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: US 60/284,962  
; PRIOR FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: US 60/286,645

; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: US 60/346,986  
; PRIOR FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 113  
; LENGTH: 645  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-126-103-113

Query Match 43.8%; Score 1872; DB 14; Length 645;  
Best Local Similarity 55.8%; Pred. No. 9e-144;  
Matches 359; Conservative 97; Mismatches 137; Indels 50; Gaps 11;

QY 1 MWIQTMTGROQTHVDSLSRLTKVELRRKIQELHVEPGLQRLFRYRGKQMEDGHTLFD 60  
Db 1 MWIQTMTGROQTHVDSLSRLTKVELRRKIQELHVEPGLQRLFRYRGKQMEDGHTLFD 60  
QY 61 YEVRLNDTIQLLRQSL-VLPHSTKRDSELSLSDTSGCCLGQSDSKSSTHGEAAETDS 119  
Db 61 YDVLNDIIQLLRPDHLPFGTSTQIEAK-----PCSNSPPKVKAPRVGSPNQPSTS 114  
QY 120 RPADEDMWDETEGLYKVNVEVDARDTNMGAWFEAQQVVRVTR-----KAPSRD---- 167  
Db 115 ARA---RLIDPGFGIYKVNELVDARDVGLGAWFEAHIHVSSTRASDQSGRGKTPLNKSSC 171  
QY 168 -----EPCSSSRP---ALBEDVIYHVKYDDYPNGVVMQNSRDVRA 206  
Db 172 KRTNGNIKHSKENTNKLDSPVSTNSDCVAADSDVIYHIQYDEYPSGTLNEMVVKDLRP 231  
QY 207 RARTIIKQWQLEVGQVVMNPNPNKXERGFWDYDAEISR-KRETRTARELYANVVLG--D 263  
Db 232 RARTILKNELNVGDVVMVNVNVPSPGQGFWDYDAEITTLKTISRTKKELRVKIFLGSE 291  
QY 264 DSLNDCRIIFVDEYFKIERPGESEPM--VDNPMRKSPGSKCHKDDVNLRCVCAHLG 321  
Db 292 GTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLERNDPECDLGGDPEKKHSCSRVC 350  
QY 322 GGRQDPKQMLCDMCDMAFIYCLDPLSSVPSDEWYCPCEKNDASVVLGELRLRESK 381  
Db 351 GKGHEPNMQLCDECNVAIHYICLNPDLKVPBEEYWCPSCKTDSSEVVKAGERLKWSK 410  
QY 382 KNAKMASATSSQORDWKGMACVGRTECTIVPSNHYGPIPGIPVGTWRRFRVQVSEGV 441  
Db 411 KKAQPSASTESRRDWRGMACVGRTECTIVPSNHYGPIPGIPVGTWRRFRVQVSEAGV 470  
QY 442 HRPHVAGHGRSNDGYSILVLAGYEDVDHGNFTYTGSGGRDLSGNKRKTAEOQCDOKL 501  
Db 471 HRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSAQTL 530  
QY 502 TMTNPAALNCFAPINDQEGAEKDWRSKGPVVRVNVNKGKNSKYAPAEAGNRYDGIYKV 561  
Db 531 TMTNPAALNCDAPLDDKIGAESRNWRACKPVRVIRSFKGRKISKYAPAEAGNRYDGIYKV 590  
QY 562 VKYWEKGS-KFLVWRYLLRDRDDDEPGWTKGKDRIKKGL 603  
Db 591 VKYWEISSHGHFLVWRYLLRDRDDVEPAPWTSEGIERSRRLCL 633

RESULT 6  
US-10-431-096-113  
; Sequence 113, Application US/10431096  
; Publication No. US20040086896A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB  
; FILE REFERENCE: D0108A CIP  
; CURRENT APPLICATION NUMBER: US/10/431,096  
; CURRENT FILING DATE: 2003-05-07  
; PRIOR APPLICATION NUMBER: US 60/284,962  
; PRIOR FILING DATE: 2001-04-19



Matches	180;	Conservative	66;	Mismatches	161;	Indels	233;	Gaps	21;
---------	------	--------------	-----	------------	------	--------	------	------	-----

[illegible]

RESULT 11  
 US-10-424-599-174378  
 ; Sequence 174378, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kowalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules  
 ; TITLE OF INVENTION: plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223) B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 174378  
 ; LENGTH: 709  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_128482C.1.pcp  
 US-10-424-599-174378

Query Match	14.7%;	Score 626.5;	DB 15;	Length 709;
Best Local Similarity	28.1%;	Pred. No. 5.6e-42;		
Matches 190: Conservative	66;	Mismatches 161;	Indels 233;	Gaps 21

```

QY      613  EALANRERKSKRE  628
        |||||||
DB      121  EALANRERKSKRE  136

RESULT 9
US-09-764-864-1301
; Sequence 1301, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1301
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (93)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1301

```

```
Query Match      15.7%; Score 672; DB 9; Length 133;
Best Local Similarity 99.2%; Pred. No. 9.8e-47;
Matches 126; Conservative 0; Mismatches 1; Indels 0; Gaps 0
```

667	KKTKEVPYSLTAQSSLIREDKSNAKLWNEVLASLKD	PASGSPFOFLSKVEET	QCIC	726
7	KKTKEVPYSLTAQSSLIREDKSNAKLWNEVLASLKD	PASGSPFOFLSKVEET	QCIC	66
727	COELVFRPIITVCOHNVCCKDLDRSFRAQVFC	PACRYDLGRSYAMVNQPLQTVLNOLF		786
67	COELVFRPIITVCOHNVCCKDLDRSFRAQVFC	PACRYDLGRSYAMVNQPLQTVLNOLF		126
787	PGYNGNR	793		
127	PGYNGNR	133		

```

RESULT 10
US-10-425-114-39797
; Sequence 39797, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 39797
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700763470_FLI.pep
US-10-425-114-39797

```

Query Match 14.7%; Score 626.5; DB 15; Length 694;  
Best Local Similarity 28.1%; Pred. No. 5.5e-42;





APPLICANT: Ratcliffe, Oliver  
APPLICANT: Adam, Luc J  
APPLICANT: Reuber, T. Lynne  
APPLICANT: Reddie, James E  
APPLICANT: Brown, Pierre E  
APPLICANT: Pilgrim, Marsha L  
APPLICANT: Dubell III, Arnold T  
APPLICANT: Pineda, Omaira  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
FILE REFERENCE: MEI-0047 CIP  
CURRENT APPLICATION NUMBER: US/10/374,780A  
CURRENT FILING DATE: 2003-02-25  
PRIOR APPLICATION NUMBER: 09/837,944  
PRIOR FILING DATE: 2001-04-18  
PRIOR APPLICATION NUMBER: 60/310,847  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 09/934,455  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/336,049  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/338,692  
PRIOR FILING DATE: 2001-12-11  
PRIOR APPLICATION NUMBER: 10/171,468  
PRIOR FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: 10/225,066  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,067  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,068  
PRIOR FILING DATE: 2002-08-09  
NUMBER OF SEQ ID NOS: 2906  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 2314  
LENGTH: 617  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
OTHER INFORMATION: G373  
US-10-374-780A-2314

Query Match 13.8%; Score 588.5; DB 15; Length 617;  
Best Local Similarity 27.6%; Pred. No. 5.9e-39;  
Matches 174; Conservative 58; Mismatches 161; Indels 237; Gaps 22;

Qy 318 CHLCGRQDPKQLMCDCECMAPHYICLDPPLSSVSEDEWYCPBC----- 363  
Db 15 CMRCQVNPSEETLTCGTCVTPHVPCLLPE-SLASSTGEWECPCDCSGVVVPSAAPGTGN 73  
Qy 364 -RNDASEVVL-----AGELRSEKKNKAKMASATSSQRP----- 396  
Db 74 ARPSSGSLVAIRAIAQADETLTAEAKKRQKLMGSGGDDGVDEEKKLEIFCSICI 133  
Qy 397 -----WGKG-----MAC-----VG 405  
Db 134 QLPERPITPCGHNFCLCFERKAVGQGLTCMI CRSKIPRHVAKNPRINLALVSAIRLA 193  
Qy 406 RTKECTI-----VPSNHYGPIP- 422  
Db 194 NVTKCSVEATAKVHHI IRNQDRPEKAFITERRAVTGTAKANAASGKPFVITPRDHFQPIA 253  
Qy 423 -----GIPVGTMMFRVQVSESGVHRPHVAGIHRNSDSYSLVLAGGYEDDVHCGN 474  
Db 254 ENDVTRKQGLVNGESWEDRQCRQWGAHPFHITAGIAGQSAVGAQSAVSGVGGVDDDDHGE 313  
Qy 475 FFFYTGSGGRDLGSKNR-TAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPV 533  
Db 314 WFDYTGSGGRDLSGNKRNKNSDQA FKNMESLSC-----KMGYPV 358  
Qy 534 RVVRNVKGGKNSKYAPAEGRYDGIYKVKVWPEKSG-FLVWRVLLRRDDDEPGPWTK 592  
Db 359 RVVRNWK-EKRSAYAPAEGRYDGVVRIEKCMSNVGVQSGFKVCRYLFRCDNEPAPWTS 417

Qy 593 -EGKRIKKLGLTMOYPEGVLEALANRERENKSKSEEBEEOQGGFASPRGTGKWKWKKS 651  
Db 418 DEHGDRPRL---PNTPE--LETAADLFVRKESPSWDFDEA-----GRWKWKWK-- 461  
Qy 652 AGGSPSRAGSPRRTSKTKTVEPYSLTAQOSSLIREDKSNAKLWNEVLASLKDRPASGSPF 711  
Db 462 -----SP-----PVSRRMALDP---EERKKKRAKNTMKARLL----- 490  
Qy 712 QLFSLKVEETPQCICQELVFRPITTVCOHNVCKCLDRSF-----R 753  
Db 491 -----KEFSQICREVLSLSEVTTTCANFNKACLEAKFAGITQLRERSNGRKLRAK 542  
Qy 754 AQVFSPACRYDLGRSYA-MQVNQPLQTVL 782  
Db 543 KNMTCPCTTDLSEFLQNPQVNRMEI 572

RESULT 14  
US-10-437-963-156872  
; Sequence 156872, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 156872  
; LENGTH: 750  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_56499C.1.pap  
US-10-437-963-156872

Query Match 11.7%; Score 498; DB 16; Length 750;  
Best Local Similarity 24.6%; Pred. No. 1.9e-31;  
Matches 167; Conservative 63; Mismatches 162; Indels 288; Gaps 25;

Qy 294 MERKSGPSCKHKOD-VNRLCRVCACHLCGRQDPKQLMCDCECMAPHYICLD--PPLS 350  
Db 1 MASSSNPNSLPCSSDGVCMCKVLTTEV-----EQLRCSTCATPWHPTCLSSIPPLT 52  
Qy 351 SVPSDEWYCPBCRND-----SEVV----- 371  
Db 53 DVA---HWVPCDGDVTASYPSPDVVRPESPESLIAIRVIEADPVLSTQEKARRQELLG 109  
Qy 372 -----LAGRLRESKKNKAKMASATSSQ-----R 395  
Db 110 HAGDAGAAITEAVGENVEDSESNFLMLNKNINCSFOMLLPERPVTTPCGHNFCLCKPR 169  
Qy 396 DW---GKGMACV-----GRTECTI----- 412  
Db 170 RNIENGK-RACVICRAPHITQKVAQDLIRINLALVQAIRMAKANAANASTTGETTVYHYKNE 228  
Qy 413 -----VPSNHYGPI-----PGIPVGTMMFRVQ 435  
Db 229 DKPDRAFTTERAKRAGMANASSGQIFVTIAPDFGPILEDHDPNRNRGVRVGDHWKDRME 288  
Qy 436 VSESGVHRPHVAGIHRNSDSYSLVLAGGYEDDVHGNFFYTGSGGRDLSGNKRTA-E 494  
Db 289 GRQWGAFFHIAQAGSTHGAQSAVSGVGGVDDDEHGEWFLYTGSGGRDLSGNKRTSKE 348  
Qy 495 QSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVKGGKNSKYAPAEGRN 554







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:41:42 ; Search time 27 Seconds

(without alignments)  
1947.784 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 4272

Sequence: 1 MM1QVMTDGRQTHTVDSLS.....VNQPLQTVLNQLPPGYGNR 793

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/aaa/5A COMB.pap:\*  
2: /cgn2\_6/ptodata/1/aaa/5B COMB.pap:\*  
3: /cgn2\_6/ptodata/1/aaa/6A COMB.pap:\*  
4: /cgn2\_6/ptodata/1/aaa/6B COMB.pap:\*  
5: /cgn2\_6/ptodata/1/aaa/PCTUS COMB.pap:\*  
6: /cgn2\_6/ptodata/1/aaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	177	4.1	4019	US-09-854-133-425	Sequence 425, App
2	169.5	4.0	1674	US-09-418-710-1	Sequence 1, Appl
3	169.5	4.0	1674	US-09-839-479-1	Sequence 1, Appl
4	167.5	3.9	1673	US-09-418-710-70	Sequence 70, Appl
5	167.5	3.9	1673	US-09-839-479-69	Sequence 69, Appl
6	158.5	3.7	2289	US-09-051-019-2	Sequence 2, Appl
7	158	3.7	371	US-09-233-342A-5	Sequence 5, Appl
8	154	3.6	391	US-09-233-342A-2	Sequence 2, Appl
9	152.5	3.6	405	US-08-881-857-2	Sequence 2, Appl
10	152.5	3.6	405	US-09-233-342A-2	Sequence 2, Appl
11	152	3.6	800	US-09-270-767-45282	Sequence 4282, A
12	151.5	3.5	1722	US-09-538-092-1033	Sequence 1033, Ap
13	148	3.5	351	US-07-945-295-2	Sequence 2, Appl
14	148	3.5	351	PCT-US91-06418-1	Sequence 2, Appl
15	147	3.4	1527	US-09-418-710-27	Sequence 27, Appl
16	147	3.4	1527	US-09-839-479-27	Sequence 27, Appl
17	147	3.4	1531	US-09-418-710-29	Sequence 29, Appl
18	147	3.4	1531	US-09-839-479-29	Sequence 29, Appl
19	145.5	3.4	1912	US-08-913-832A-2	Sequence 2, Appl
20	145.5	3.4	1912	US-09-249-181A-2	Sequence 2, Appl
21	145.5	3.4	1912	US-09-158-707-2	Sequence 2, Appl
22	144.5	3.4	1121	US-07-789-915A-2	Sequence 2, Appl
23	144.5	3.4	1121	US-08-005-002C-2	Sequence 2, Appl
24	144.5	3.4	1121	US-08-487-203A-2	Sequence 2, Appl
25	144	3.4	504	US-09-270-767-45920	Sequence 45920, A
26	144	3.4	1969	US-09-418-710-72	Sequence 72, Appl
27	144	3.4	1969	US-09-839-479-71	Sequence 71, Appl

28	144	3.4	1972	4	US-09-418-710-21	Sequence 21, Appl
29	144	3.4	1972	4	US-09-839-479-21	Sequence 21, Appl
30	143	3.3	1525	4	US-09-418-710-69	Sequence 69, Appl
31	143	3.3	1525	4	US-09-839-479-68	Sequence 68, Appl
32	139.5	3.3	1242	4	US-09-488-270A-2	Sequence 44, Appl
33	138.5	3.2	45	4	US-09-839-479-44	Sequence 2, Appl
34	135.5	3.2	352	3	US-08-854-764-2	Sequence 2, Appl
35	135.5	3.2	352	5	PCT-US95-09377-2	Sequence 43, Appl
36	134.5	3.1	45	4	US-09-418-710-43	Sequence 56, Appl
37	134.5	3.1	45	4	US-09-418-710-56	Sequence 42, Appl
38	134.5	3.1	45	4	US-09-839-479-42	Sequence 55, Appl
39	134.5	3.1	45	4	US-09-839-479-55	Sequence 45, Appl
40	131.5	3.1	45	4	US-09-418-710-45	Patent No. 5510474
41	131.5	3.1	533	6	5510474-2	Sequence 58, Appl
42	131	3.1	44	4	US-09-418-710-58	Sequence 57, Appl
43	131	3.1	44	4	US-09-839-479-57	Sequence 7, Appl
44	128.5	3.0	156	2	US-09-070-060-7	Sequence 3, Appl
45	128.5	3.0	156	3	US-09-051-969A-3	

## ALIGNMENTS

RESULT 1  
US-09-854-133-425  
; Sequence 425, Application US/09854133  
; Patent No. 6759508  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.475C10  
; CURRENT APPLICATION NUMBER: US/09/854,133  
; CURRENT FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 735  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 425  
; LENGTH: 4019  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-854-133-425

Query Match	4.1%	Score 177	DB 4	Length 4019
Best Local Similarity	21.6%	Pred. No. 3.2e-06		
Matches 109	Conservative 58	Mismatches 150	Indels 188	Gaps 25
Qy	313	CRVC-ACHLCGRDDPKQLMCDSCDMAFHLYCLDPLLSVPSSEDEW---YCPERN-DA	367	
Db	59	CLETCVCAAGKATDPORLLICDDCDISYHYCLDPLQTPV-KGGWKCKWVWCRHGA	117	
Qy	368	SEVVLAGEIRLESKNKAKMASATS-----SSQDWMGKMACVGRTKCTTVPSPNHYGIP	422	
Db	118	TSAGLRCEWQNNYTQCAPCASLSCPCVYRNEEDLILQC-----ROC-----	161	
Qy	423	GIPVTWRRFV-----QVSESG-----VHRPHVAGIHGRSNDGYSVLVLAG--G	465	
Db	162	-----DRMHAVCONLTETEENEVADIGDCSMCPYMPASNPVSSDCESSIQAIVT	216	
Qy	466	YEDDVHGNFPTVTG-----SGGRDLSG-----NKRTEQSCDQKLTNTNRALNCFA	514	
Db	217	KVLEDPKTYTODGCLTESGTMQTSITVTPRRKSKPKLKLIINQNSAVLQTPP	276	
Qy	515	PINDQEG--AAAKWRSKGPVRVVRNVKGGKSKYAPAEGRNRYDGIYKVYKWPKEKSG	572	
Db	277	DIQSEHSRDEGMDDSREGE-----LMDCDKSESS-----PEREAV--	312	
Qy	573	FLVWRVLLRRDDDEPGPWTKGDKRIKKLGLTMQYPEGYLEALANREKENSKEEEQ	632	
Db	313	-----DDETKG---VEGTDGVKK-----RKRKPYRP	335	

QY 512 CFAPINDOGAEAKDWRGKP-VRVVRNVKGGKNSKYAPAGNRYDGYIKVVKVYPEK-- 568  
 1371 -----QEEBVSILPKGRPOVRLPVKTRGKLSFSSFSRGOQBPGRY-----PSRSQ 1417  
 QY 569 -----GKSGFLVWRYLLRDDDEPFWTKGDKRIKLLTWMQYPEG----- 610  
 1418 QSTPKTVSSKTGRS-----LRKINSAPPTETKS-----LRASRSTRSHSGPLOADV 1465  
 QY 611 YLEALANRERKENKREEEEOEGGFASPRCKGKWKRSAGGSPS-----RAGSPRR 664  
 1466 FVELLSPRKRGRKRSANNTPENSPNFRVIAIKSQSRSVNTASKLSLOESSEKRR 1525  
 QY 665 TSKTKVPEYSLTAQOOS-----LIREDKSNAKLWNEVLASLKDRP 705  
 1526 CRKRQSPSPVTLGRSSRGQGVHLSAFEQVLVVELVRHDDS-----W----- 1570  
 QY 706 ASGSPFQFLSKVETFCICCOELVFRPI-----TTVCOHNVCKDCLD 749  
 1571 -----PFLKLVSKI-----QVPDYVDIIKPIALNIIREKVNKCEYKLASEFID 1614  
 Db  
 RESULT 3  
 US-09-839-479-1  
 ; Sequence 1, Application US/09839479  
 ; Patent No. 6727222  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jones, Michael H.  
 ; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR  
 ; FILE REFERENCE: 06501-042002  
 ; CURRENT APPLICATION NUMBER: US/09/839,479  
 ; CURRENT FILING DATE: 2001-04-20  
 ; PRIOR APPLICATION NUMBER: US 09/418,710  
 ; PRIOR FILING DATE: 1999-10-15  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/01783  
 ; PRIOR FILING DATE: 1998-04-17  
 ; PRIOR APPLICATION NUMBER: JP 9/310027  
 ; PRIOR FILING DATE: 1997-10-24  
 ; PRIOR APPLICATION NUMBER: JP 9/116570  
 ; PRIOR FILING DATE: 1997-04-18  
 ; NUMBER OF SEQ ID NOS: 72  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1674  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-839-479-1  
 Query Match 4.0%; Score 169.5; DB 4; Length 1674;  
 Best Local Similarity 18.3%; Pred. No. 3.9e-06;  
 Matches 141; Conservative 100; Mismatches 244; Indels 287; Gaps 35;  
 QY 81 HSTKERDSELSDT-----DSGCCLGQSESDKSTHGEAAATDSRPA-----DED 125  
 1027 NSRGHRESALKETLLQEKSRICAQLARFSEKHFSDKPDQPSKPTYSRGRSSNAYDPSQ 1086  
 QY 126 MWDETELGLYKNEVVDART-----NMGAWFEAQVVRTRKAPSRDEPCSSSTRPALED 181  
 1087 MCAEQLELRDLFLDIEDRIYQGTG-----IKVTDR-----HIWRSALSEG 1131  
 QY 182 VIYHKYDDYPNGVVMNSRDV-----ARARTIIKWQDL-----E 218  
 1132 -RYELLSSENGKGIKTVNEDEVEEMEIDEQTKVIVKDLGLIKTETPTSTVSTNASTPOS 1190  
 QY 219 VQGVV-----MLNYPNDPKRGFW---YDAEISRKRETRTARELYANVVLGDDSLNDCRI 271  
 1191 VSSVHYLAMALFQIEQGIERRFLKAPLDASDS-GRSYKTVLDRWRESLSSASLSQ--- 1246  
 QY 272 IFVDEVKIERGEGSPMVDNPMRRKSGPSCKHCKDDVNLRCVACHLCGGRQDPDKOL 331  
 1247 VFL-HLSTLDR-----SVIWSKILNARCKICR-----KKGDAENNV 1282  
 QY 332 MCDECDMAFIYCLDPLSSVPSDEWYCPCECNDAEVLVLAGERLRESKKNAMASATS 391

QY 633 QEGGF---ASPTGKGKWKRSAGGSPSRAGSPRRTSKTKKVEPYSLTAQOQSLIREDKS 689  
 336 GIGGFVWRORSTGQKTKR-----SVIRKDSG 363  
 QY 690 NA-----KLWNEVLA-SLKDRPASGSP-----FQFLSKVETFCICCOELV 731  
 364 GSISEQLPCDDGWSQPLDVLDESVSVTESTEKIKRYRKRKNLEBTFPAY-LQEAF 422  
 QY 732 FRPITTVCOHNVCKDCLDRSFRAQV 756  
 423 FG-----KDLIDTSROSKI 436  
 Db  
 RESULT 2  
 US-09-418-710-1  
 ; Sequence 1, Application US/09418710  
 ; Patent No. 6596482  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jones, Michael H.  
 ; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR  
 ; FILE REFERENCE: 06501-042001  
 ; CURRENT APPLICATION NUMBER: US/09/418,710  
 ; CURRENT FILING DATE: 1999-10-15  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/01783  
 ; PRIOR FILING DATE: 1998-04-17  
 ; PRIOR APPLICATION NUMBER: JP 9/310027  
 ; PRIOR FILING DATE: 1997-10-24  
 ; PRIOR APPLICATION NUMBER: JP 9/116570  
 ; PRIOR FILING DATE: 1997-04-18  
 ; NUMBER OF SEQ ID NOS: 73  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1674  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-418-710-1  
 Query Match 4.0%; Score 169.5; DB 4; Length 1674;  
 Best Local Similarity 18.3%; Pred. No. 3.9e-06;  
 Matches 141; Conservative 100; Mismatches 244; Indels 287; Gaps 35;  
 QY 81 HSTKERDSELSDT-----DSGCCLGQSESDKSTHGEAAATDSRPA-----DED 125  
 1027 NSRGHRESALKETLLQEKSRICAQLARFSEKHFSDKPDQPSKPTYSRGRSSNAYDPSQ 1086  
 QY 126 MWDETELGLYKNEVVDART-----NMGAWFEAQVVRTRKAPSRDEPCSSSTRPALED 181  
 1087 MCAEQLELRDLFLDIEDRIYQGTG-----IKVTDR-----HIWRSALSEG 1131  
 QY 182 VIYHKYDDYPNGVVMNSRDV-----ARARTIIKWQDL-----E 218  
 1132 -RYELLSSENGKGIKTVNEDEVEEMEIDEQTKVIVKDLGLIKTETPTSTVSTNASTPOS 1190  
 QY 219 VQGVV-----MLNYPNDPKRGFW---YDAEISRKRETRTARELYANVVLGDDSLNDCRI 271  
 1191 VSSVHYLAMALFQIEQGIERRFLKAPLDASDS-GRSYKTVLDRWRESLSSASLSQ--- 1246  
 QY 272 IFVDEVKIERGEGSPMVDNPMRRKSGPSCKHCKDDVNLRCVACHLCGGRQDPDKOL 331  
 1247 VFL-HLSTLDR-----SVIWSKILNARCKICR-----KKGDAENNV 1282  
 QY 332 MCDECDMAFIYCLDPLSSVPSDEWYCPCECNDAEVLVLAGERLRESKKNAMASATS 391  
 1283 LCDGDRGHHTYCVRPKLATVP-EGDWFCPECK-----PKQRCRLS--- 1323  
 QY 392 SSQRDWGKGMACVGRTEKCTIVPSNIHYGPIGTPVGTWMEFRVQVSESGVHREHVAGIHG 451  
 1324 -----FR-----QRESLES--- 1332  
 QY 452 RSDNGSVSLVAGGYEDDVGHNFFVTYTGSGGRDLSGNKRTABQSCDQKLTNTNRAALN 511  
 1333 -DEDVEDSM---GGEDDEVD-----GDEEGQSEEBEYVEQDEDDDS----- 1370  
 Db

```
Db 1283 LCDGDRGHHTYCVRPKLKTP-EGDWFCEPCR- 1323
QY 392 SSORDWKGKMACVGTREKCTIVPSNHGYPGIPVGTWVRFRVQVSGVHRPHVAGIHG 451
Db 1324 -----FR-----QPSLES- 1332
QY 452 RSDGYSVLVAGGYDDVDHGNFFTYTSGGRDLGNKRTAEQSCDQKLTNTNRLALN 511
Db 1333 -DEDVEDSM--GGEDEVD-----GDEEGQSEEEVEVEQDEDD- 1370
QY 512 CFAPINDOEGAEAKWRSGKP-VRVVRNVKGGKNSKYAPAEAGNRYDGIYKVYVWPEK-- 568
Db 1371 -----QEEVEVSLPKRGPQVRLPVKTRGKLSLSSFSRSGQQCEPGRY-----PERSQ 1417
QY 569 -----GKSGFLVWRYLLRRDDDEPGWTKEGDKRIKGLGTMQYPEG----- 610
Db 1418 QSTPKTTVSSKTRGS-----LRKINSAPPTETKS-----LRIASRSTRHSHGPLQADV 1465
QY 611 YLEALANREREKENSKEEEEOEGGFASPRGKWKWKSAGGSPS-----RAGSPRR 664
Db 1466 FVELLSPRKRRGRKSANNTPENSPNFRVIATKSSSEGRSVNIASKLSLOESSEKRR 1525
QY 665 TSKTKTVPYSLTAQSS-----LIREKSNAKLMNEVLASLKDRP 705
Db 1526 CRKQSPSPVTLGRSSRGQGVHLSAFEQVLVVELVRHDS--W----- 1570
QY 706 ASGSPFLSKVEETFOCIQCOELVFRPI-----TTVCOHNVCKDCLD 749
Db 1571 -----PFLKLVSKI-----QVPDYDIKKPIALNIIREKVNKCEYKLASEFID 1614

RESULT 4
US-09-418-710-70
; Sequence 70, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-042001
; CURRENT APPLICATION NUMBER: US/09/418,710
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01783
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 1673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-710-70

Query Match 3.9%; Score 167.5; DB 4; Length 1673;
Best Local Similarity 18.3%; Pred. No. 6.1e-06;
Matches 141; Conservative 99; Mismatches 245; Indels 287; Gaps 35;

QY 81 HSTKERDSELSDT-----DSGCCLGQSESDKSSSTHGEAAETDSRPA-----DED 125
Db 1026 NSGRHRESALKETLLQEKSRICAQLARFSEKHFSDKRPDPSKPTYSRGRSSNAYDPSQ 1085
QY 126 MWDETELGLKVNVEYDART-----NMGAMFEAQVVRVTKAPSRDEPCSTSRPALED 181
Db 1086 MCAEQLELRFLDIEDRIIVQGTUGA-----IKVTD-----HIWRSALSG 1130
QY 182 VIHVYKYDDPENGVMQNSRDVR-----APARTIIKWQDL-----E 218
Db 1131 -RYELLSEENKENGIIKTVDNEVEIMEIDQTVIVKORLLGHTTPTSTVNASTPQS 1189
QY 219 VGOVV-----MLNYPNDPNKRGFW-----YDAEISRKRETRTARELYANVLGDSDLNDCRI 271
```

```
Db 1190 VSSVHYLAMALFQEQGLERRFLKAPLDASDS--GRSKTVLDRWRESLLSASLSQ--- 1245
QY 272 IFDDEVFKIERPCGSGPMWNPMPKSGPCKCHCKDDNRLCRVCAACHLCGRGDDPKQL 331
Db 1246 VFL-HLSTLOR-----SVIWSKSLNARCKICR-----KKGDAENMV 1281
QY 332 MCDECDMAPIHYCLDPPSLSPVSEDEWYCPCECRNDASEVLAGERLRRESKNAKMASATS 391
Db 1282 LCDGDRGHHTYCVRPKLKTP-EGDWFCEPCR- 1322
QY 392 SSORDWKGKMACVGTREKCTIVPSNHGYPGIPVGTWVRFRVQVSGVHRPHVAGIHG 451
Db 1323 -----FR-----QPSLES- 1331
QY 452 RSDGYSVLVAGGYDDVDHGNFFTYTSGGRDLGNKRTAEQSCDQKLTNTNRLALN 511
Db 1332 -DEDVEDSM--GGEDEVD-----GDEEGQSEEEVEVEQDEDD- 1369
QY 512 CFAPINDOEGAEAKWRSGKP-VRVVRNVKGGKNSKYAPAEAGNRYDGIYKVYVWPEK-- 568
Db 1370 -----QEEVEVSLPKRGPQVRLPVKTRGKLSLSSFSRSGQQCEPGRY-----PERSQ 1416
QY 569 -----GKSGFLVWRYLLRRDDDEPGWTKEGDKRIKGLGTMQYPEG----- 610
Db 1417 QSTPKTTVSSKTRGS-----LRKINSAPPTETKS-----LRIASRSTRHSHGPLQADV 1464
QY 611 YLEALANREREKENSKEEEEOEGGFASPRGKWKWKSAGGSPS-----RAGSPRR 664
Db 1465 FVELLSPRKRRGRKSANNTPENSPNFRVIATKSSSEGRSVNIASKLSLOESSEKRR 1524
QY 665 TSKTKTVPYSLTAQSS-----LIREKSNAKLMNEVLASLKDRP 705
Db 1525 CRKQSPSPVTLGRSSRGQGVHLSAFEQVLVVELVRHDS--W----- 1569
QY 706 ASGSPFLSKVEETFOCIQCOELVFRPI-----TTVCOHNVCKDCLD 749
Db 1570 -----PFLKLVSKI-----QVPDYDIKKPIALNIIREKVNKCEYKLASEFID 1613

RESULT 5
US-09-839-479-69
; Sequence 69, Application US/09839479
; Patent No. 6727222
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-042002
; CURRENT APPLICATION NUMBER: US/09/839,479
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 09/418,710
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01783
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 1673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-839-479-69

Query Match 3.9%; Score 167.5; DB 4; Length 1673;
Best Local Similarity 18.3%; Pred. No. 6.1e-06;
Matches 141; Conservative 99; Mismatches 245; Indels 287; Gaps 35;

QY 81 HSTKERDSELSDT-----DSGCCLGQSESDKSSSTHGEAAETDSRPA-----DED 125
Db 1026 NSGRHRESALKETLLQEKSRICAQLARFSEKHFSDKRPDPSKPTYSRGRSSNAYDPSQ 1085
```

CLASSIFICATION:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2289 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-09-051-019-2

Query Match 3.7%; Score 158.5; DB 3; Length 2289;  
Best Local Similarity 21.5%; Pred. No. 7.3e-05;  
Matches 79; Conservative 49; Mismatches 154; Indels 85; Gaps 14;

126 MWDETELGKLVKNEVVDADT-----NMGNWFAQVVRVTRKAPSRDEPCSTSRPALBED 181  
1086 MCAEQKLELRDLFDIEDRIYQGTGLG-----IKVTR-----HWRSALBSG 1130  
182 VYHVVKYDDYPNGVVMNSRVR-----ARARTIKWQDL-----E 218  
1131 RYELLSEENKENGILKTIVEDVEEIDEQKIVKORLLGKLTETPTSTVSTNASTPOS 1189  
219 VQGVV-----MLNYPNDPNKRGFW-----YDAEISRKRETRTARELYANVVLGDDSLNDCRI 271  
1190 VSSVHYLAVALFOIEQGLERFLKAPLDASDS-GRSYKTVLDRWRESLLSSASUSQ----- 1245  
272 IFVDEVFKIERPCEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVACACHLCGGRRQDPDKOL 331  
1246 VFL-HLSTLDR-----SVIWSKSIILNARCKICR-----KKGDAENMV 1281  
332 MCDECDMAFHICLDPLSPVSEDEWYPCPCRNDAASEVVLAGERLRESKKNAKMASATS 391  
1282 LCDGCRGHHTYCVRPKLIKVP-EGDWFCPCR-----PKQRCRLS--- 1322  
392 SSQDMGKGMACVGRTEKCTIVSNHYGPICIPVGTMTWFRVQVSESGVHRPHVAGIHG 451  
1323 -----FR-----QPSLES--- 1331  
452 RSNDSYSLVLAGYEDVDHGNFFVTYSGGRDLGSKNKTAEQSCDQKLTNTNRLALN 511  
1332 -DEDVEDSM--GGEDDEV-----GDEEGQSEEEVEVEQDEDS--- 1369  
512 CFAPINDQGAEAADMRSKGP-VRVVRNVKGGKSNKYAPAEGRNYDGIYKVKVWPEK-- 568  
1370 -----QBEESVSLPKRGPOVELPVKTRGKLSSTSPSRQOQEPGRY-----PSRSQ 1416  
569 -----GKSGFLWRYLLRRDDDEPGWTKGKORIKKLGLTMOYPEG----- 610  
1417 QSTPKTVSSKTGRS-----LRKINSAPPTTKS-----LRASRSTRSHGLOADV 1464  
611 YLEALANRERKENSKEEBEQGGFASPTGKWKWKRSAGGSPS-----RAGSPRR 664  
1465 FVELLSPPRRKRRGRKSNANTPENSFPNFRVIATKSSQESQSVNIAKSLQSESESKRR 1524  
665 TSCKTKVPYSLTAQSS-----LIREDKSNAKLWNEVLASLKDPR 705  
1525 CEKQSPSPVTLGRSSRGQGVHLSAFEQLVVELVRHDS-----W----- 1569  
706 ASGSPFQLKSVETFOCICQELVFRPI-----TTVCOHNVCKDCLD 749  
1570 -----PFLKLVSKI-----QVPDYDIKKPEALNIIRKKNKCEYKLASEFID 1613

RESULT 6  
US-09-051-019-2  
; Sequence 2, Application US/09051019  
; Patent No. 6103229  
; GENERAL INFORMATION:  
; APPLICANT: KAHMANN, Regine and QUADBECK-SEEGER, Claudia  
; TITLE OF INVENTION: Regulatory gene from Ustilago maydis  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Keil & Weinlauf  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage  
; OPERATING SYSTEM: IBM AT-compatible, Pentium processor  
; SOFTWARE: WordPerfect version 6.1  
; CURRENT APPLICATION DATA: US/09/051,019  
; FILING DATE: 31-MAR-1998

US-09-233-342A-5  
; Sequence 5, Application US/09233342A  
; Patent No. 6207803  
; GENERAL INFORMATION:  
; APPLICANT: KIKLY, KRISTINE K.  
; APPLICANT: GROSS, MITCHELL S.  
; APPLICANT: HURLE, MARK ROBERT  
; TITLE OF INVENTION: HUMAN REQUIEM  
; FILE REFERENCE: ATG-50013-1  
; CURRENT APPLICATION NUMBER: US/09/233,342A  
; CURRENT FILING DATE: 1999-01-19  
; PRIOR APPLICATION NUMBER: 08/881,857  
; PRIOR FILING DATE: 1997-06-24  
; PRIOR APPLICATION NUMBER: 60/021,299  
; PRIOR FILING DATE: 1996-06-26  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-233-342A-5

Query Match 3.7%; Score 158; DB 3; Length 371;  
Best Local Similarity 19.8%; Pred. No. 4.1e-06;  
Matches 82; Conservative 46; Mismatches 123; Indels 164; Gaps 19;

52 MEDGHTLFDYEVRLNDTIQLLVRSQSLVLPSTHSTKERSELSDTSGCCLQSESDKSTHG 111  
1 MEQCH-----NYNARLCAERSVRLPF-----LDSQTGVAQSNVCYINWKEHRGP--G 45



Matches 80; Conservative 39; Mismatches 128; Indels 140; Gaps 16;

QY 15 TVDSLSRATK-----VSELRKIQELPHVEPLQRLFYR---GKQMEDGHTLFDYEVRL 65  
Db 71 TLDSQEDVTEDFVSIIPWIRERIIDL---EANIERRYLYKPLGSGTGDH-LAVIAQNG 126

QY 66 NDTIQ-----LVQRQSLVLPSTKRSSELSDTDSGCCCLGQSDSKSTHGEAA 114  
Db 127 HTTQTQNSASAAVILQWQQQQQQQQQQQQQQGAGNSLNP-SFNERTMALAA 185

QY 115 AETDSRPADENWDETELGLYKNEVVDARDTNMGAFEAQVVRVTRKAPSRDEPCSSTS 174  
Db 186 AAAASGPGNA-----TGVAASAVAGATPCESGSGSEPNSGNA 222

QY 175 RPAAEEDVIYHKYDDYPENGVMNSRDVRRARARTIIRKQDLEVGQVVMNLNPNPK 234  
Db 223 SPASNCSDRDEKVEQIPK-GLVQ----- 245

QY 235 RGFWDYDAEISRRKET-RTARELYANVVLGDDSLNDRIRIIFVDFVKIERGEGSPWVDP 293  
Db 246 ---WRDA-VSRSHTTAQLAMALYV-----LESC-VAMDKSIMKAN----- 280

QY 294 MRRKSGPSCKHCKDDVNRCLRCVCACHLGGGRQDPDKOLMCDCEMDMAFHYICLDPPLSSVP 353  
Db 281 -----COPCTSGENEDKLLLCDCGCKGHTYCFKPKMDNIP 316

QY 354 SEDEWYCPCRNDAE-----VVLAGER 376  
Db 317 DGD-WYCYECVKNATNERKCIIVCGHR 342

RESULT 12

US-09-538-092-1033  
; Sequence 1033, Application US/09538092  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuratSeqformatter Version 0.9  
; SEQ ID NO 1033  
; LENGTH: 1722  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number P29375

US-09-538-092-1033

Query Match 3.5%; Score 151.5; DB 4; Length 1722;  
Best Local Similarity 25.7%; Pred. No. 0.00021;  
Matches 53; Conservative 25; Mismatches 63; Indels 65; Gaps 8;

QY 191 YP-----ENGV-----VQMSNRDVRARARTIIRKQDLEVGQVVMNLNPNPKRGFWYDAE 242  
Db 167 YFVELFQSGVSLMGVQMPNLDKEV-----EPEVLSTDTQTSPEFGTMMNLL 214

QY 243 ISRKETRTARELYANVVLGDDSLNDRIRIIFVDFVKIERGEGSPWVDPNPKRGSGPSC 302  
Db 215 PKRTRRVKTQSE-----SGDVSRN-----TELKQLIFGAGPKVGLAM-----GT 255

QY 303 KHCKDDVNRCLRCV----- 376  
Db 256 KKEDEVTRRRKVTNRSDAFNQMQRKGTLSVNFVDLYVCMFCGRGNEDKLLLCDCGCD 315

QY 312 LCRVCACHLGGGRQDPDKOLMCDCEMDMAFHYICLDPPLSSVSEDEWYCPEC 363  
Db 343 -----CCNICGTSENDDQLLFCDCCDRGYHMYCLTPSMSE-PPEGWSWSCHLC 388

RESULT 10

US-09-233-342A-2  
; Sequence 2, Application US/09233342A  
; Patent No. 6207803  
; GENERAL INFORMATION:  
; APPLICANT: KIKUY, KRISTINE K.  
; APPLICANT: GROSS, MITCHELL S.  
; APPLICANT: HURLE, MARK ROBERT  
; TITLE OF INVENTION: HUMAN REQUIEM  
; FILE REFERENCE: ATG-50013-1  
; CURRENT APPLICATION NUMBER: US/09/233,342A  
; PRIOR FILING DATE: 1999-01-19  
; PRIOR APPLICATION NUMBER: 08/881,857  
; PRIOR FILING DATE: 1997-06-24  
; PRIOR APPLICATION NUMBER: 60/021,299  
; PRIOR FILING DATE: 1996-06-26  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS

US-09-233-342A-2

Query Match 3.6%; Score 152.5; DB 3; Length 405;  
Best Local Similarity 22.0%; Pred. No. 1.6e-05;  
Matches 51; Conservative 27; Mismatches 63; Indels 91; Gaps 10;

QY 160 TRKAPSRDEPCSSTSRPAAEEDVIYHKYDDYPENGVMNSRDVRRARARTIIRKQDLEV 219  
Db 220 TSKAPQRCVGRKYNRPEGLS-----YHVAHSLAE-----EGEDKED 257

QY 220 GOVVMNLNPNPKRGFWYDAEISRRKETRTARE-----LYANVVLGDDSLNDR 270  
Db 258 SQ-----PTTP-----VSORSEQKSKKGPDPGLALPNNYCDCLGDSKIN--- 297

QY 271 IIFVDFVKIERGEGSPWDP-NPMWRKSGPSCK-----HCKDDVNR 311  
Db 298 -----KKTGOPELVSCDCGRSGHPSCLOFTPVMAAVKTYRWQCIECK----- 342

QY 312 LCRVCACHLGGGRQDPDKOLMCDCEMDMAFHYICLDPPLSSVSEDEWYCPEC 363  
Db 343 -----CCNICGTSENDDQLLFCDCCDRGYHMYCLTPSMSE-PPEGWSWSCHLC 388

RESULT 11

US-09-270-767-45282  
; Sequence 45282, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45282  
; LENGTH: 800  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-45282

Query Match 3.6%; Score 152; DB 4; Length 800;  
Best Local Similarity 20.7%; Pred. No. 5.3e-05;



QY 338 MAFHIYCLDPLSLVSPSEDEWYCPEC 363  
 Db 316 DSYHTFCLIPLEDVPKGD-WRCPKC 340

## RESULT 13

US-07-945-295-2  
 ; Sequence 2, Application US/07945295  
 ; Patent No. 6610823  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Targoff, Ira N.  
 ; APPLICANT: Ge, Qun  
 ; TITLE OF INVENTION: Antigens Associated with Polymyositis  
 ; TITLE OF INVENTION: and with Dermatomyositis  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Kilpatrick & Cody  
 ; STREET: 1100 Peachtree Street, Ste. 2800  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: United States  
 ; ZIP: 30309-4530  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: IBM PC compatible  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/945,295  
 ; FILING DATE: 19920909  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pabst, Patrea L.  
 ; REGISTRATION NUMBER: 31,284  
 ; REFERENCE/DOCKET NUMBER: OMRP 120CIP  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (404) 815-6508  
 ; TELEFAX: (404) 815-6558  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 351 amino acids  
 ; TYPE: AMINO ACID  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FRAGMENT TYPE: N-terminal  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapiens  
 ; TISSUE TYPE: Blood  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: Human thymocyte lambda gt11  
 ; CLONE: L1  
 ; US-07-945-295-2

Query Match 3.5%; Score 148; DB 4; Length 351;  
 Best Local Similarity 24.9%; Pred. No. 3.3e-05;  
 Matches 48; Conservative 22; Mismatches 61; Indels 62; Gaps 8;  
 QY 216 DLEVGQVVMNPNPNKRGFWYDAISRKRETRTARE-LYANVVLGDDSLNDCRIIFV 274  
 Db 29 DLDSGSHSASGRPDGPVR-----TKLKGPRGPKKKKVLGCPAVAGEEVDGYETHQ 83  
 QY 275 DEVFKIERPGE-----GSPVMDNPMRRKSGP-SCKHC----- 305  
 Db 84 DYCEVCOQGGIEILDCPCPRAYHLVCLDPELD---RAPEGKWSCPHCKEKGQVQWEAKEEE 140  
 QY 306 -----KDDVNRLCRVACACILCGGRQDPDKOLMCDECDMAFHIYCLDPPLS 350  
 Db 141 EYEEEGEERGEKEEDDHMEYCRVC-----KGGELLCDDACTISSYHIHCLNPPLP 192

QY 351 SVPSDEWYCPEC 363  
 Db 193 DIPN-GEWLCPRC 204

RESULT 14  
 PCT-US91-06418-1  
 ; Sequence 1, Application PC/TUS9106418  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Oklahoma Medical Research, Foundation, et al  
 ; TITLE OF INVENTION: Antigens Associated with Polymyositis  
 ; TITLE OF INVENTION: and with Dermatomyositis  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Kilpatrick & Cody  
 ; STREET: 100 Peachtree Street  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: US  
 ; ZIP: 30303  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US91/06418  
 ; FILING DATE: 19910905  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/579023  
 ; FILING DATE: 09-JUL-1990  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pabst, Patrea L.  
 ; REGISTRATION NUMBER: 31,284  
 ; REFERENCE/DOCKET NUMBER: OMRF120  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 404-572-6508  
 ; TELEFAX: 404-572-6555  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 351 amino acids  
 ; TYPE: AMINO ACID  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FRAGMENT TYPE: N-terminal  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapien  
 ; TISSUE TYPE: Sera  
 ; IMMEDIATE SOURCE:  
 ; CLONE: L1  
 ; PCT-US91-06418-1

Query Match 3.5%; Score 148; DB 5; Length 351;  
 Best Local Similarity 24.9%; Pred. No. 3.3e-05;  
 Matches 48; Conservative 22; Mismatches 61; Indels 62; Gaps 8;  
 QY 216 DLEVGQVVMNPNPNKRGFWYDAISRKRETRTARE-LYANVVLGDDSLNDCRIIFV 274  
 Db 29 DLDSGSHSASGRPDGPVR-----TKLKGPRGPKKKKVLGCPAVAGEEVDGYETHQ 83  
 QY 275 DEVFKIERPGE-----GSPVMDNPMRRKSGP-SCKHC----- 305  
 Db 84 DYCEVCOQGGIEILDCPCPRAYHLVCLDPELD---RAPEGKWSCPHCKEKGQVQWEAKEEE 140  
 QY 306 -----KDDVNRLCRVACACILCGGRQDPDKOLMCDECDMAFHIYCLDPPLS 350  
 Db 141 EYEEEGEERGEKEEDDHMEYCRVC-----KGGELLCDDACTISSYHIHCLNPPLP 192  
 QY 351 SVPSDEWYCPEC 363

us-10-019-071-2.ra1

Tue Nov 2 11:31:39 2004

Db 193 DIPN-GEWLCPRC 204

```

RESULT 15
US-09-418-710-27
; Sequence 27, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-042001
; CURRENT APPLICATION NUMBER: US/09/418,710
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01783
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-710-27

```

```

Query Match 3.4%; Score 147; DB 4; Length 1527;
Best Local Similarity 36.7%; Pred. No. 0.00046;
Matches 29; Conservative 13; Mismatches 29; Indels 8; Gaps 2;

QY 318 CHLGGGRQDPKQMLMCDCEMAMHICYICLDPPLSSVPSSEDEWYCPICRNDASEVVLGERL 377
Db 1187 CKVCPKGGDDKLLCDECNKAFHLFCLRPALYVVP-DGEWQCPACQP-----ATARR 1238

QY 378 RESKKNAMASATSSQSD 396
Db 1239 NSRGNYTEESASEDSED 1257

```

Search completed: November 1, 2004, 15:45:47  
Job time : 30 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:41:42 ; Search time 91 Seconds  
(without alignments)  
3126.070 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 4272

Sequence: 1 MVIQRTMDGRQTHTVDSLS.....VNPQLTVLNQLPPGYGNGR 793

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4272	100.0	793	4	Abb76980 Human Inv
2	4263	99.8	793	6	Abc48157 Human bla
3	4263	99.8	793	7	Adf61820 Human NP9
4	4263	99.8	793	7	Adf61820 Human NP9
5	4263	99.8	793	8	Adf61820 Human NP9
6	4263	99.8	793	8	Adn05229 Antipsori
7	4191	98.1	780	6	Ado20357 Human PRO
8	4191	98.1	780	6	Abu56628 Lung canc
9	2233.5	53.7	802	6	Adn38838 Cancer/an
10	1872	43.8	645	6	Abc82238 Human nuc
11	1872	43.8	645	6	Abu69599 Human NF-
12	926	21.7	174	4	Abu69620 Human NF-
13	823.5	19.3	198	5	Abb76983 Human Inv
14	672	15.7	133	4	Abb64013 Human ORF
15	672	15.7	133	6	Abu16348 Human nov
16	644	15.1	233	3	Abu55417 Human nov
17	588.5	13.8	617	7	Abd43114 Human ORF
18	588.5	13.8	617	8	Adi43851 Plant yie
19	519	12.1	150	4	Adm19888 Protein e
20	471	11.0	148	4	Adm20128 Protein e
21	420.5	9.8	178	4	Abu13350 Human nov
22	420.5	9.8	178	6	Abu55419 Human nov
23	420.5	9.8	180	4	Abu42002 Human nov
24	357.5	8.4	110	4	Abu15895 Human nov
25	357.5	8.4	110	6	Abu54964 Human nov

26	287	6.7	71	4	AA40216	Aam40216 Human pol
27	287	6.7	71	4	AA40216	Aab73737 Ring fing
28	265.5	6.2	781	3	AAG32044	Aag32044 Arabidops
29	265.5	6.2	785	3	AAG32043	Aag32043 Arabidops
30	265.5	6.2	788	3	AAG32042	Aag32042 Arabidops
31	247	5.8	89	4	AAO07182	Aao07182 Human pol
32	235.5	5.5	670	8	ADN72155	Adn72155 Thale cre
33	235.5	5.5	1072	3	AAG41664	Aag41664 Arabidops
34	235.5	5.5	1079	3	AAG41663	Aag41663 Arabidops
35	235.5	5.5	1132	3	AAG41662	Aag41662 Arabidops
36	229.5	5.4	856	3	ADM48182	Adm48182 Polypepti
37	193.5	4.5	1398	5	ABG97491	Abg97491 Human NOV
38	193.5	4.5	1400	4	AA833348	Aab83348 AAP-2 pro
39	193.5	4.5	1445	6	ABR41365	Abc41365 Human sec
40	190	4.4	87	5	AB897911	Abb97911 Human sec
41	187	4.4	2176	8	ADN73253	Adn73253 Thale cre
42	185	4.3	399	8	ADP22650	Adp22650 Sea-squir
43	180.5	4.2	2897	4	ABB58514	Abb58514 Drosophil
44	180	4.2	884	4	ABB71641	Abb71641 Drosophil
45	177.5	4.2	4952	5	ADH47759	Adh47759 NOV18 pro

## ALIGNMENTS

RESULT 1  
ABB76980  
ID ABB76980 standard; protein; 793 AA.  
XX  
AC ABB76980;  
XX  
DT 22-JUL-2002 (first entry)  
XX  
DE Human Inverted CCAAT box binding protein, ICBP90.  
XX  
KW Human; inverted CCAAT box binding protein; ICBP90; cytostatic;  
XX  
KW cell proliferation control; inverted CCAAT box; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200078949-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-FR001747.  
XX  
PR 22-JUN-1999; 99FR-00007935.  
XX  
PS (ADDER-) ADEREGEM ASSOC DEV RECH EN GENETIQ.  
XX  
PI Bronner C, Hopfner R, Mousli M, Jeltsch J, Lutz Y, Oudet P;  
DR WPI; 2001-091571/10.  
XX  
PS N-PSDB; ABL58020.  
XX  
FT Novel inverted CCAAT box binding protein, and related nucleic acids,  
XX  
FT antibodies and specific ligands, useful for treating and preventing  
XX  
FT cancer.  
XX  
PS Claim 1; Fig 7; 115pp; French.  
XX  
CC The present sequence is the protein sequence for human ICBP90 (inverted  
XX  
CC CCAAT box binding protein). The inverted CCAAT box is implicated in cell  
XX  
CC proliferation control. Several copies of the inverted CCAAT box are  
XX  
CC present in the promoter of the topoisomerase IIalpha gene, and also  
XX  
CC functions as a nuclear receptor. ICBP90 and its coding sequence are  
XX  
CC useful for treatment and/or prevention of cancer  
XX  
SQ Sequence 793 AA;

Query Match 100.0%; Score 4272; DB 4; Length 793;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

16-JAN-2003.  
03-JUL-2002; 2002WO-US021338.  
03-JUL-2001; 2001US-0302814P.  
03-AUG-2001; 2001US-0310099P.  
08-NOV-2001; 2001US-0343705P.  
13-NOV-2001; 2001US-0350666P.  
12-APR-2002; 2002US-0372246P.  
(EOSB-) EOS BIOTECHNOLOGY INC.  
Mack DH, Aziz N;  
WPI; 2003-201532/19.  
N-PSDB; ACC50965.  
Detecting a bladder cancer-associated transcript in a cell from a patient, comprises contacting a biological sample from the patient with a bladder cancer-associated polynucleotide or antibody.  
Claim 10; Page 238; 307pp; English.  
The present invention describes a method for detecting a bladder cancer-associated transcript in a cell from a patient. The method comprises contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is 80 % identical to a table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059 encode the human bladder cancer-associated proteins given in ABR48146 to ABR48421. Bladder cancer-associated sequences from the present invention have cytostatic activities, and can be used in antisense gene therapy and in vaccine production. The method can be used for detecting a bladder cancer-associated transcript in a cell from a patient. The method is useful in diagnosing or treating bladder cancer and in screening for compounds that modulate bladder cancer, such as hormones or antibodies. The nucleic acid molecules from the present invention may be used in various screening and diagnostic methods, and for gene therapy, vaccine and/or antisense/inhibition applications  
Sequence 793 AA;  
Query Match 99.8%; Score 4263; DB 6; Length 793;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 791; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MWIQVRTMDGROTHVTVDLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60  
Db 1 MWIQVRTMDGROTHVTVDLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60  
QY 61 YEVRINDTIQLLVRSQSLVPHSTKERDSELSDTSGCCLGQSGESDKSTHGEAAETDSR 120  
Db 61 YEVRINDTIQLLVRSQSLVPHSTKERDSELSDTSGCCLGQSGESDKSTHGEAAETDSR 120  
QY 121 PADEDMWDETELGLYKNEYVDARTNMGANPEAQVVRTRKAPSEDEPCSTSRPALEE 180  
Db 121 PADEDMWDETELGLYKNEYVDARTNMGANPEAQVVRTRKAPSEDEPCSTSRPALEE 180  
QY 181 DVIYHVKYDDYPENGVOQNSRDVBARARTIIKWQDLEVGQVVMNPNPKERGFWD 240  
Db 181 DVIYHVKYDDYPENGVOQNSRDVBARARTIIKWQDLEVGQVVMNPNPKERGFWD 240  
QY 241 AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGEPSGMVDMPMRKSGP 300  
Db 241 AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGEPSGMVDMPMRKSGP 300  
QY 301 SCCHKDDVNRLCRVCACHLCGGRODPDKQLMCDCECDMAFHLYCLDPLSSVSEDEWYC 360  
Db 301 SCCHKDDVNRLCRVCACHLCGGRODPDKQLMCDCECDMAFHLYCLDPLSSVSEDEWYC 360  
QY 361 PECRNDASEVVLAGEFLRESKKNKAKWASATSSQORDWKGMACVGRKTECTIIVPSNHYGP 420  
Db 361 PECRNDASEVVLAGEFLRESKKNKAKWASATSSQORDWKGMACVGRKTECTIIVPSNHYGP 420

1 MWIQVRTMDGROTHVTVDLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60  
1 MWIQVRTMDGROTHVTVDLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60  
61 YEVRINDTIQLLVRSQSLVPHSTKERDSELSDTSGCCLGQSGESDKSTHGEAAETDSR 120  
61 YEVRINDTIQLLVRSQSLVPHSTKERDSELSDTSGCCLGQSGESDKSTHGEAAETDSR 120  
121 PADEDMWDETELGLYKNEYVDARTNMGANPEAQVVRTRKAPSEDEPCSTSRPALEE 180  
121 PADEDMWDETELGLYKNEYVDARTNMGANPEAQVVRTRKAPSEDEPCSTSRPALEE 180  
181 DVIYHVKYDDYPENGVOQNSRDVBARARTIIKWQDLEVGQVVMNPNPKERGFWD 240  
181 DVIYHVKYDDYPENGVOQNSRDVBARARTIIKWQDLEVGQVVMNPNPKERGFWD 240  
241 AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGEPSGMVDMPMRKSGP 300  
241 AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERPGEPSGMVDMPMRKSGP 300  
301 SCCHKDDVNRLCRVCACHLCGGRODPDKQLMCDCECDMAFHLYCLDPLSSVSEDEWYC 360  
301 SCCHKDDVNRLCRVCACHLCGGRODPDKQLMCDCECDMAFHLYCLDPLSSVSEDEWYC 360  
361 PECRNDASEVVLAGEFLRESKKNKAKWASATSSQORDWKGMACVGRKTECTIIVPSNHYGP 420  
361 PECRNDASEVVLAGEFLRESKKNKAKWASATSSQORDWKGMACVGRKTECTIIVPSNHYGP 420  
421 IPGIPVGTWFRVQVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVGHNFFYTG 480  
421 IPGIPVGTWFRVQVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVGHNFFYTG 480  
481 SGGDLISGNKRTAGSCDOKLTNTNRLALNCFAPINDQEGAEAKDWRSGKPVVVRNVK 540  
481 SGGDLISGNKRTAGSCDOKLTNTNRLALNCFAPINDQEGAEAKDWRSGKPVVVRNVK 540  
541 GGNKSKYAPAEGRNYDGIYKVYKWEKSGFLVWRYLLRRDDDEPGFWTKEGDKRIKK 600  
541 GGNKSKYAPAEGRNYDGIYKVYKWEKSGFLVWRYLLRRDDDEPGFWTKEGDKRIKK 600  
601 LGLTQWPEGYLEALANRERKENSKRREEEQEGGFASPTGKGKWRKSAGGSPSRAG 660  
601 LGLTQWPEGYLEALANRERKENSKRREEEQEGGFASPTGKGKWRKSAGGSPSRAG 660  
661 SPRRTSKTKVPEYSLTAQGSLSIREDKSNKLNNEVLASLKDRLPASGSPQLFLSKVEE 720  
661 SPRRTSKTKVPEYSLTAQGSLSIREDKSNKLNNEVLASLKDRLPASGSPQLFLSKVEE 720  
721 TFQCICCOELVFRPITTVCOHNVCKDCLDRSPRAQVFCPACRYDLGRSYAMQVNOPLQT 780  
721 TFQCICCOELVFRPITTVCOHNVCKDCLDRSPRAQVFCPACRYDLGRSYAMQVNOPLQT 780  
781 VINQLFPYGYNGR 793  
781 VINQLFPYGYNGR 793  
RESULT 2  
ABR48157  
ID ABR48157 standard; protein; 793 AA.  
XX AC ABR48157;  
XX AC ABR48157;  
DT 12-JUN-2003 (first entry)  
XX Human bladder cancer associated protein sequence SEQ ID NO:27.  
DE Human; bladder cancer; cytostatic; gene therapy; vaccine.  
KW Human; bladder cancer; cytostatic; gene therapy; vaccine.  
OS Homo sapiens.  
XX WO2003003906-A2.  
EN

QY 421 IPGIPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSYLVLAGGYEDDVGHNFFTYTG 480  
DB 421 IPGIPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSYLVLAGGYEDDVGHNFFTYTG 480  
QY 481 SGRDLGSKNRTAEQSCDQKLTNTNRALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540  
DB 481 SGRDLGSKNRTAEQSCDQKLTNTNRALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540  
QY 541 GGNKSKYAPAEGRNYDGIYKVVKYWEKSGSGLVWRYLLRRDDDEPGPWTKGDKRIKK 600  
DB 541 GGNKSKYAPAEGRNYDGIYKVVKYWEKSGSGLVWRYLLRRDDDEPGPWTKGDKRIKK 600  
QY 601 LGLTMOYPEGYLEALANRERKENSKEEEOEGGFASPRTGKWKRSAGGGSFRAG 660  
DB 601 LGLTMOYPEGYLEALANRERKENSKEEEOEGGFASPRTGKWKRSAGGGSFRAG 660  
QY 661 SPRRTSKTKVPEYSLTAQSSLIREDKSNKLNWNEVLASLKDORPASGSPFOLFSLKVEE 720  
DB 661 SPRRTSKTKVPEYSLTAQSSLIREDKSNKLNWNEVLASLKDORPASGSPFOLFSLKVEE 720  
QY 721 TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRQAQVFCPCRYDLGRSYAMQVNOPLQT 780  
DB 721 TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRQAQVFCPCRYDLGRSYAMQVNOPLQT 780  
QY 781 VLNLQFPFGYNGR 793  
DB 781 VLNLQFPFGYNGR 793

## RESULT 3

ID ADF61820  
ADP61820 standard; protein; 793 AA.

AC ADP61820;

DT 12-FEB-2004 (first entry)  
DE Human NP95 protein.

KW cell cycle arrest; cytostatic; antiapoptotic; antiarteriosclerotic;  
KW vasotrophic; antichyroid; melanoma; breast; ovarian; lung;  
KW gastrointestinal; colon cancer; Grave's disease; psoriasis;  
KW atherosclerosis; restenosis; vasoproliferative; human; NP95.

OS Homo sapiens.

XX WO2003088910-A2.  
XX 30-OCT-2003.

XX 15-APR-2003; 2003WO-US011867.  
XX 15-APR-2002; 2002US-00123568.

XX 15-APR-2002; 2002US-00123731.  
XX 16-APR-2002; 2002US-0373366P.

XX (RIGE-) RIGEL PHARM INC.

XX Hitoshi Y. Jenkins Y;

XX WPI; 2003-865396/80.

XX DR N-PSDB; ADF61819.

PT Identifying a compound that modulates cell cycle arrest, for treating  
PT e.g. cancer, comprises contacting a cell comprising a target polypeptide  
PT and determining the chemical or phenotypic effect of the compound upon  
PT the cell.  
XX Claim 1; SEQ ID NO 4; 176pp; English.

XX The invention relates to a novel method for identifying a compound that

CC modulates cell cycle arrest comprising contacting a target polypeptide

CC within a cell with a compound and determining the chemical or phenotypic  
CC effect of the compound upon the cell. The method of the invention has  
CC cytostatic, antiapoptotic, antiarteriosclerotic, vasotrophic and  
CC antichyroid applications and may be useful for identifying a compound  
CC that modulates cell cycle arrest. Such compounds may subsequently be used  
CC for developing therapeutic reagents to treat melanoma, breast, ovarian,  
CC lung, gastrointestinal or colon cancer, as well as other proliferative  
CC diseases such as Grave's disease, psoriasis, atherosclerosis, restenosis  
CC and other vasoproliferative diseases. The current sequence is that of the  
CC human NP95 protein of the invention.

XX Sequence 793 AA;

Query Match 99.8%; Score 4263; DB 7; Length 793;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 791; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWIQTVMGROTHVTDSLSRLTKVEELRRKIQELPHVPEGLQRLFYRGKQMEDGHTLFD 60

DB 1 MWIQTVMGROTHVTDSLSRLTKVEELRRKIQELPHVPEGLQRLFYRGKQMEDGHTLFD 60

QY 61 YEVRNDTIQLLVROSLVLPSTKERDSELSPTDSCCLGQSESDKSTHGEAAETDSR 120

DB 61 YEVRNDTIQLLVROSLVLPSTKERDSELSPTDSCCLGQSESDKSTHGEAAETDSR 120

QY 121 PADEDMWDETELGLYKVNVEYVDARTNMGAWFEAQVVRVTRKAPSRDEPCSTSRPALEE 180

DB 121 PADEDMWDETELGLYKVNVEYVDARTNMGAWFEAQVVRVTRKAPSRDEPCSTSRPALEE 180

QY 181 DVIYHVKKDDYPENGWQNSRDVAPARTIIKWQDLEVGQVVMNLNPNPKRGFWYD 240

DB 181 DVIYHVKKDDYPENGWQNSRDVAPARTIIKWQDLEVGQVVMNLNPNPKRGFWYD 240

QY 241 AEISKRTRTARELYANVWLDDSLNDCRIIFVDEVEKIERPBGSGSWNDPMRKSGP 300

DB 241 AEISKRTRTARELYANVWLDDSLNDCRIIFVDEVEKIERPBGSGSWNDPMRKSGP 300

QY 301 SKCHKDDVNLRCVCAHLCCGRQDDPKQMLCDECDMAFIYCLDPLSSVPSDEWYC 360

DB 301 SKCHKDDVNLRCVCAHLCCGRQDDPKQMLCDECDMAFIYCLDPLSSVPSDEWYC 360

QY 361 PECDNDASEVVLAGERLRESKNAKASATSSQORDWKGMACVGRTKECTIVPSNHYP 420

DB 361 PECDNDASEVVLAGERLRESKNAKASATSSQORDWKGMACVGRTKECTIVPSNHYP 420

QY 421 IPGIPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSYLVLAGGYEDDVGHNFFTYTG 480

DB 421 IPGIPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSYLVLAGGYEDDVGHNFFTYTG 480

QY 481 SGRDLGSKNRTAEQSCDQKLTNTNRALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540

DB 481 SGRDLGSKNRTAEQSCDQKLTNTNRALNCFAPINDQEGAEAKDWRSGKPVVRNVK 540

QY 541 GGNKSKYAPAEGRNYDGIYKVVKYWEKSGSGLVWRYLLRRDDDEPGPWTKGDKRIKK 600

DB 541 GGNKSKYAPAEGRNYDGIYKVVKYWEKSGSGLVWRYLLRRDDDEPGPWTKGDKRIKK 600

QY 601 LGLTMOYPEGYLEALANRERKENSKEEEOEGGFASPRTGKWKRSAGGGSFRAG 660

DB 601 LGLTMOYPEGYLEALANRERKENSKEEEOEGGFASPRTGKWKRSAGGGSFRAG 660

QY 661 SPRRTSKTKVPEYSLTAQSSLIREDKSNKLNWNEVLASLKDORPASGSPFOLFSLKVEE 720

DB 661 SPRRTSKTKVPEYSLTAQSSLIREDKSNKLNWNEVLASLKDORPASGSPFOLFSLKVEE 720

QY 721 TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRQAQVFCPCRYDLGRSYAMQVNOPLQT 780

DB 721 TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRQAQVFCPCRYDLGRSYAMQVNOPLQT 780

QY 781 VLNLQFPFGYNGR 793

DB 781 VLNLQFPFGYNGR 793

RESULT 4	
AD76781	
ID	AD76781 standard; protein; 793 AA.
XX	
AC	AD76781;
XX	
DT	26-FEB-2004 (first entry)
XX	
DE	Novel human secreted and transmembrane protein SeqID 456.
XX	
KW	human; PRO; membrane bound protein; membrane bound receptor;
KW	cell proliferation; cell migration; cell differentiation;
KW	mitogenic factor; survival factor; cytotoxic factor;
KW	differentiation factor; neuroepitide; hormone; cell receptor;
KW	receptor-ligand interaction; cytoskeletal; chondrocyte; tumour.
XX	
OS	Homo sapiens.
XX	
PN	WO2003072035-A2.
XX	
PD	04-SEP-2003.
XX	
PF	21-FEB-2003; 2003WO-US005241.
XX	
PR	22-FEB-2002; 2002US-0359461P.
XX	
PA	(GETH ) GENENTECH INC.
XX	
PI	Bedary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;
PI	Williams PM, Wood WI, Wu TD;
XX	
DR	WPI: 2003-721702/68.
DR	N-PSDB; AD76780.
XX	
PT	New PRO polypeptides, useful for diagnosing and treating an immune
PT	related disorder, e.g. systemic lupus erythematosus, rheumatoid
PT	arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
PT	diabetes mellitus.
XX	
PS	Claim 10; SEQ ID NO 456; 918pp; English.
XX	
CC	This invention relates to novel nucleic acids encoding human PRO secreted
CC	and transmembrane proteins. Extracellular proteins play important roles
CC	in the formation, differentiation and maintenance of multicellular
CC	organisms. The fate of many individual cells (for example proliferation,
CC	migration or differentiation) is typically governed by information
CC	received from other cells and the immediate environment. The information
CC	is often transmitted by secreted polypeptides (for example mitogenic
CC	factors, survival factors, cytotoxic factors, differentiation factors,
CC	neuropeptides and hormones) which are received and interpreted by diverse
CC	cell receptors or membrane bound proteins. These membrane bound proteins
CC	and receptors may be of use as pharmaceutical and diagnostic agents, such
CC	as in the blocking of receptor-ligand interactions. The current invention
CC	provides the amino acid sequences of novel human membrane bound receptors
CC	and proteins, along with the cDNA sequences encoding them. The novel
CC	proteins of the invention may have cytostatic activities through the
CC	stimulation of chondrocytes. The nucleic acids of the invention may be
CC	useful for the manufacture of a medicament for diagnosing or treating a
CC	tumour in a mammal. In addition, they may be useful for measuring or
CC	detecting the expression of a tumour associated gene. The present
CC	invention.
XX	
SQ	Sequence 793 AA;
Query Match	99.8%; Score 4263; DB 7; Length 793;
Best Local Similarity	99.7%; Pred. No. 0;
Matches 791; Conservative	1; Mismatches 1; Indels 0; Gaps 0;
QY	1 MWIQRVTMDGRQTHVTVDLSRLTKVEELRRKIKQLFHFVPEGLQRLFYRGKQMEDGHTLPD 60
DB	1 MWIQRVTMDGRQTHVTVDLSRLTKVEELRRKIKQLFHFVPEGLQRLFYRGKQMEDGHTLPD 60

QY	61	YEVRLNDTIQLLAVROSLLVPHSTKERSLSLSDTSCCLGQSESDKSTHGEAAETDSR	120
DB	61	YEVRLNDTIQLLAVROSLLVPHSTKERSLSLSDTSCCLGQSESDKSTHGEAAETDSR	120
QY	121	PADEDMWDETELGLYKVNVEYVDARDTNMGAFEAQVVRTRKAPSEDEPCSSTSRPALEE	180
DB	121	PADEDMWDETELGLYKVNVEYVDARDTNMGAFEAQVVRTRKAPSEDEPCSSTSRPALEE	180
QY	181	DVIYHVKYDDYPENGVVQNSRDVRARATIIKWQLEVGQVVMANVNDNPKERGFWD	240
DB	181	DVIYHVKYDDYPENGVVQNSRDVRARATIIKWQLEVGQVVMANVNDNPKERGFWD	240
QY	241	AEISRKRETRARELYANVVLGDDSLNDCRIIFVDEVFKIERPEGEGSPMDNPMRRKSGP	300
DB	241	AEISRKRETRARELYANVVLGDDSLNDCRIIFVDEVFKIERPEGEGSPMDNPMRRKSGP	300
QY	301	SKHKCKDDVNRLCRVCACHLCGRQDPDKQLMCDCECDMAFHIIYCLDPLSSVSEDESWYC	360
DB	301	SKHKCKDDVNRLCRVCACHLCGRQDPDKQLMCDCECDMAFHIIYCLDPLSSVSEDESWYC	360
QY	361	PECRNDASEVVLAGERLRESKKNAKMASATSSSSORDMGKGMACVGRTECTIIVPSNHYGP	420
DB	361	PECRNDASEVVLAGERLRESKKNAKMASATSSSSORDMGKGMACVGRTECTIIVPSNHYGP	420
QY	421	IPGIPVGTMMRFVQVSESGVHRPHVAGIHGRSNDGSYSILVLAGGYEDDVGHNFFTYTG	480
DB	421	IPGIPVGTMMRFVQVSESGVHRPHVAGIHGRSNDGSYSILVLAGGYEDDVGHNFFTYTG	480
QY	481	SGGRDLGSGNKRRTAEQSCDQKLTNTNRALALNCFAPINDOGEAEAKOWRSGKPVVRNVK	540
DB	481	SGGRDLGSGNKRRTAEQSCDQKLTNTNRALALNCFAPINDOGEAEAKOWRSGKPVVRNVK	540
QY	541	GGKNSKYAPAEGRNYDGIYKVKYKVPKSGFLWRYLLRRDDDEPGMTKEGDKRIKK	600
DB	541	GGKNSKYAPAEGRNYDGIYKVKYKVPKSGFLWRYLLRRDDDEPGMTKEGDKRIKK	600
QY	601	LGLTMQYPEGVLEALANREKENSKEEEBQOEGGFASPRGTGKWKRSAGGPGSRAG	660
DB	601	LGLTMQYPEGVLEALANREKENSKEEEBQOEGGFASPRGTGKWKRSAGGPGSRAG	660
QY	661	SPRTSKTKTVEPYSLTAQSSLIREDKSNAKLWNEVLASLKORPASGSPQLFLSKVEE	720
DB	661	SPRTSKTKTVEPYSLTAQSSLIREDKSNAKLWNEVLASLKORPASGSPQLFLSKVEE	720
QY	721	TFQCICQELVFRDITVTCQHNVCCKDCIDRSFRAQVFCPCACRYDGLGRSYAMQVNPLOT	780
DB	721	TFQCICQELVFRDITVTCQHNVCCKDCIDRSFRAQVFCPCACRYDGLGRSYAMQVNPLOT	780
QY	781	VLNQLFFGYGNR 793	
DB	781	VLNQLFFGYGNR 793	
RESULT 5			
ADN05229			
ID	ADN05229	standard; protein; 793 AA.	
XX			
AC	ADN05229;		
XX			
DT	01-JUL-2004	(first entry)	
XX			
DE	Antipsoriatic protein sequence #791.		
XX			
KW	antipsoriatic; gene therapy; psoriasis; diagnosis.		
OS	Homo sapiens.		
XX			
PN	WO2004028479-A2.		
XX			
PD	08-APR-2004.		
XX			
PF	25-SEP-2003; 2003WO-US030907.		

XX PR 25-SEP-2002; 2002US-0414006P.  
XX PA (GETH ) GENENTECH INC.  
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WT;  
XX PI Wu TD;  
XX DR WPI: 2004-305105/28.  
XX DR N-PSDB; ADN05228.  
XX PT New PRO nucleic acid or polypeptide, useful for preparing a  
XX PT pharmaceutical composition for diagnosing or treating psoriasis in a  
XX PT mammal.  
XX PS Claim 9; SEQ ID NO 1623; 3069pp; English.  
XX CC The invention relates to novel polynucleotide and polypeptides for  
XX CC treating psoriasis or a sequence having at least 80% identity to the  
XX CC above sequences. The nucleic acid is useful for preparing a composition  
XX CC for diagnosing or treating psoriasis in a mammal. This sequence  
XX CC corresponds to one of the polypeptides of the invention.  
XX SQ Sequence 793 AA;  
Query Match 99.8%; Score 4263; DB 8; Length 793;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 791; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MWIQRMTDGRQTHTVDSLSRLTKVEELRRKIQELFHFVEFGLQRLFYRKQKQEDGHTLFD 60  
DB 1 MWIQRMTDGRQTHTVDSLSRLTKVEELRRKIQELFHFVEFGLQRLFYRKQKQEDGHTLFD 60  
QY 61 YEVRNDTIQLVRSQSLVPHSTKERDSLSLTDSCCLGQSESDKSTHGEAAFTDSR 120  
DB 61 YEVRNDTIQLVRSQSLVPHSTKERDSLSLTDSCCLGQSESDKSTHGEAAFTDSR 120  
QY 121 PADEDMWDETEGLVKNEYVDARTNMGAWFEAQQVVRTRKAPSRDEPCSTSPALBE 180  
DB 121 PADEDMWDETEGLVKNEYVDARTNMGAWFEAQQVVRTRKAPSRDEPCSTSPALBE 180  
QY 181 DVIYHKVDYDPENGQVQNSRDVRAARTIITKQDLEVGQVVMNLNPDNPKERGFWD 240  
DB 181 DVIYHKVDYDPENGQVQNSRDVRAARTIITKQDLEVGQVVMNLNPDNPKERGFWD 240  
QY 241 AEISKRTRTARELYANVLGDDSLNDCRIIFVDEVFKIERPGSGPMVNDPMRKSGP 300  
DB 241 AEISKRTRTARELYANVLGDDSLNDCRIIFVDEVFKIERPGSGPMVNDPMRKSGP 300  
QY 301 SCKHCKDDVNLRCVCAHLCGRQDPDKQLMCDCECDMAFHLYCLDPLSSVPSSEDEWYC 360  
DB 301 SCKHCKDDVNLRCVCAHLCGRQDPDKQLMCDCECDMAFHLYCLDPLSSVPSSEDEWYC 360  
QY 361 PECDNDASEVVLGRLRESKNAKMASATSSQSDWKGMAQVGRTEKCTIVPSNHYGP 420  
DB 361 PECDNDASEVVLGRLRESKNAKMASATSSQSDWKGMAQVGRTEKCTIVPSNHYGP 420  
QY 421 IPGIPVGTMRFRVQVSSGVHRHVAGIHGRSNDGSLVLAGYEDVDVHGNNFTYT 480  
DB 421 IPGIPVGTMRFRVQVSSGVHRHVAGIHGRSNDGSLVLAGYEDVDVHGNNFTYT 480  
QY 481 SGGRLDGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWSGKPVVRNVK 540  
DB 481 SGGRLDGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWSGKPVVRNVK 540  
QY 541 GKGNSKAPAEGRNRYDGIYKVKYWPKEKSGFLVWRYLLRRDDDEPGPWTKEGDR1KK 600  
DB 541 GKGNSKAPAEGRNRYDGIYKVKYWPKEKSGFLVWRYLLRRDDDEPGPWTKEGDR1KK 600  
QY 601 LGLTMQVPEGYLEALANREREKENSKRREEEQQGGGFASPTGKGKWRKSAGGSPSRAG 660  
DB 601 LGLTMQVPEGYLEALANREREKENSKRREEEQQGGGFASPTGKGKWRKSAGGSPSRAG 660

QY 661 SPRTSKTKVEPYSLTAQSSLIREDKSNKLNWNEVLASLKDPSAGSPFQLFSKVEE 720  
DB 661 SPRTSKTKVEPYSLTAQSSLIREDKSNKLNWNEVLASLKDPSAGSPFQLFSKVEE 720  
QY 721 TFOCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPACRYDLGRSYAMQVNPLOT 780  
DB 721 TFOCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPACRYDLGRSYAMQVNPLOT 780  
QY 781 VLNLQFPGYGNGR 793  
DB 781 VLNLQFPGYGNGR 793  
RESULT 6  
ADO20357  
ID ADO20357 standard; protein; 793 AA.  
XX ADO20357;  
XX AC ADO20357;  
XX DT 12-AUG-2004 (first entry)  
XX DE Human PRO polypeptide #626.  
XX KW Human; PRO; immune related disorder; systemic lupus erythematosus;  
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;  
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;  
KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;  
KW diabetes mellitus; renal disease; demyelinating disease;  
KW central nervous system; peripheral nervous system;  
KW demyelinating polyneuropathy; Guillain-Barre syndrome;  
KW chronic inflammatory demyelinating polyneuropathy.  
XX OS Homo sapiens.  
XX PN WO2004043361-A2.  
XX PD 27-MAY-2004.  
XX PF 06-NOV-2003; 2003WO-US035268.  
XX PR 08-NOV-2002; 2002US-0425235P.  
XX PA (GETH ) GENENTECH INC.  
XX PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;  
XX PI Wood WT, Wu TD;  
XX DR WPI: 2004-420067/39.  
XX DR N-PSDB; ADO20356.  
XX PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for  
XX PT treating an immune related disorder such as systemic lupus erythematosus,  
XX PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or  
XX PT spondyloarthropathy.  
XX PS Claim 7; SEQ ID NO 1252; 1731pp; English.  
XX CC The invention relates to human PRO polypeptides and the polynucleotides  
XX CC encoding them. The polypeptides and polynucleotides are useful for  
XX CC treating and diagnosing immune related disorders in mammals. The immune  
XX CC related disorders include systemic lupus erythematosus, rheumatoid  
XX CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic  
XX CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune  
XX CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes  
XX CC mellitus, immune-mediated renal disease, demyelinating diseases of the  
XX CC central or peripheral nervous system, demyelinating polyneuropathy,  
XX CC Guillain-Barre syndrome and chronic inflammatory demyelinating  
XX CC polyneuropathy. This sequence represents a human PRO polypeptide of the  
XX CC invention.  
XX SQ Sequence 793 AA;  
Query Match 99.8%; Score 4263; DB 8; Length 793;

Best Local Similarity 99.7%; Pred. No. 0; Matches 791; Conservative 1; Mismatches 1; Indels 0; Gaps 0;									
Qy	1	MWIVRTMDGRQTHTVDSLSRLTKVEELRRKIQLFHVPEGLQRLFYRGKQMEDGHTLFD	60						
Db	1	MWIVRTMDGRQTHTVDSLSRLTKVEELRRKIQLFHVPEGLQRLFYRGKQMEDGHTLFD	60						
Qy	61	YEVLNDTIQLLVRSQSLVLPSTKERSLSDTSGCCLGQSDSKSTHGEAAETDSR	120						
Db	61	YEVLNDTIQLLVRSQSLVLPSTKERSLSDTSGCCLGQSDSKSTHGEAAETDSR	120						
Qy	121	PAEDMDWDETELGLYKNEVVDARDTNNGANFAQVVRTRKAPSRDEPCSTSRPALEE	180						
Db	121	PAEDMDWDETELGLYKNEVVDARDTNNGANFAQVVRTRKAPSRDEPCSTSRPALEE	180						
Qy	181	DVIYHVKYDDYPENGVQMSRDVRRARTIIKWQDLEVGQVVMNPNPKERGFWD	240						
Db	181	DVIYHVKYDDYPENGVQMSRDVRRARTIIKWQDLEVGQVVMNPNPKERGFWD	240						
Qy	241	AEISRKRETRTARELYANVVLGDSLNDCLIFVDEVFKIERPCEGSPMVDNPMRKS	300						
Db	241	AEISRKRETRTARELYANVVLGDSLNDCLIFVDEVFKIERPCEGSPMVDNPMRKS	300						
Qy	301	SKHCKDDVNLRCVACHLGGRODPDKOLMCDCEDMAPHIYCLDPLSSVPSEBWC	360						
Db	301	SKHCKDDVNLRCVACHLGGRODPDKOLMCDCEDMAPHIYCLDPLSSVPSEBWC	360						
Qy	361	PECNDASEVVLAGERLRESKKNAMASATSSSQORDWKGKMACVGRTECTIVPSNHYG	420						
Db	361	PECNDASEVVLAGERLRESKKNAMASATSSSQORDWKGKMACVGRTECTIVPSNHYG	420						
Qy	421	IPGIPVGTMRFRVQVSESGVHRPHVAGIHRSDNGSYSLVLAGGYEDVDHGNFFYTG	480						
Db	421	IPGIPVGTMRFRVQVSESGVHRPHVAGIHRSDNGSYSLVLAGGYEDVDHGNFFYTG	480						
Qy	481	SGGRDLSGNKRTAQSCDKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK	540						
Db	481	SGGRDLSGNKRTAQSCDKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK	540						
Qy	541	GGKNSKYAPAGNRYDGIYKVKYKPEKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKK	600						
Db	541	GGKNSKYAPAGNRYDGIYKVKYKPEKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKK	600						
Qy	601	LGLTWQPEGLALEANRERKENSKEEERQEGGFASPTGKWKWKAGGSPSAG	660						
Db	601	LGLTWQPEGLALEANRERKENSKEEERQEGGFASPTGKWKWKAGGSPSAG	660						
Qy	661	SPRRTSKTKVPEYSLTAQSSLIREDKSNKLNWELASLKDRLPASGSPFLSKVEE	720						
Db	661	SPRRTSKTKVPEYSLTAQSSLIREDKSNKLNWELASLKDRLPASGSPFLSKVEE	720						
Qy	721	TFQCICCOELVFRPITTVCOHNVCCKDLDRFRAQVFCPCACRYDLGRSVMQVNPLOT	780						
Db	721	TFQCICCOELVFRPITTVCOHNVCCKDLDRFRAQVFCPCACRYDLGRSVMQVNPLOT	780						
Qy	781	VLNQLFPGYGNR	793						
Db	781	VLNQLFPGYGNR	793						
RESULT 7									
ABU56628									
ID	ABU56628 standard; protein; 780 AA.								
XX									
AC	ABU56628;								
XX									
DT	02-APR-2003 (first entry)								
XX									
DE	Lung cancer-associated polypeptide #221.								
XX									
KW	Lung cancer-associated polypeptide; cytostatic; emphysema;								
KW	antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;								
KW	small cell lung cancer; benign lesion; precancerous lesion; bronchitis;								

chronic obstructive pulmonary disease; hypersensitivity pneumonitis; interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

Unidentified.

WO200286443-A2.

31-OCT-2002.

18-APR-2002; 2002WO-US012476.

18-APR-2001; 2001US-0284770P.

10-MAY-2001; 2001US-0290492P.

09-NOV-2001; 2001US-0339245P.

13-NOV-2001; 2001US-0350666P.

29-NOV-2001; 2001US-0334370P.

12-APR-2002; 2002US-0372246P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Aziz N, Murray R;

WPI, 2003-093161/08.

N-PSDB; ABX76357.

Detecting a lung cancer-associated transcript in a cell from a patient for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer.

Claim 27; Page 357; 453pp; English.

The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated polypeptides of the invention

Sequence 780 AA;

Query Match 98.1%; Score 4191; DB 6; Length 780;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 778; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MWIVRTMDGRQTHTVDSLSRLTKVEELRRKIQLFHVPEGLQRLFYRGKQMEDGHTLFD 60

Db 1 MWIVRTMDGRQTHTVDSLSRLTKVEELRRKIQLFHVPEGLQRLFYRGKQMEDGHTLFD 60

Qy 61 YEVLNDTIQLLVRSQSLVLPSTKERSLSDTSGCCLGQSDSKSTHGEAAETDSR 120

Db 61 YEVLNDTIQLLVRSQSLVLPSTKERSLSDTSGCCLGQSDSKSTHGEAAETDSR 120

Qy 121 PAEDMDWDETELGLYKNEVVDARDTNNGANFAQVVRTRKAPSRDEPCSTSRPALEE 180

Db 121 PAEDMDWDETELGLYKNEVVDARDTNNGANFAQVVRTRKAPSRDEPCSTSRPALEE 180

Qy 181 DVIYHVKYDDYPENGVQMSRDVRRARTIIKWQDLEVGQVVMNPNPKERGFWD 240

Db 181 DVIYHVKYDDYPENGVQMSRDVRRARTIIKWQDLEVGQVVMNPNPKERGFWD 240





Db 361 PECDNDASVVLAGERLRESKKAKWASATSSQDRWGKMACVGTKECTIVPSNHYGP 420  
QY 421 IPGIPVGTMMRFVQVSESGVHRPHVAGIHGRSNDGSYSVLVAGGYEDVDHGNFTYTG 480  
Db 421 IPGIPVGTMMRFVQVSESGVHRPHVAGIHGRSNDGSYSVLVAGGYEDVDHGNFTYTG 480  
QY 481 SGGRLSGNKRTRAEQSCDOKLTNTNRALALNCAPINDQEGAEAKDWRSGKPVVRNVK 540  
Db 481 SGGRLSGNKRTRAEQSCDOKLTNTNRALALNCAPINDQEGAEAKDWRSGKPVVRNVK 540  
QY 541 GGNKSVAPAEGRNYDGIYKVVWPEKSGFLVWRYLLRRDDDPGPTWTEGKDRKK 600  
Db 541 GGNKSVAPAEGRNYDGIYKVVWPEKSGFLVWRYLLRRDDDPGPTWTEGKDRKK 600  
QY 601 LGITMQYPEGYLEALANRERENKREBEQEGGFASPRTGKWKRSAGGSPSRAG 660  
Db 601 LGITMQYPEGYLEALANRERENKREBEQEGGFASPRTGKWKRSAGGSPSRAG 660  
QY 661 SPRTSKTKVPEYSLTAQOSSLIREDKSNKLNWELASLKDRPASGSPQLFLSKVEE 720  
Db 661 SPRTSKTKVPEYSLTAQOSSLIREDKSNKLNWELASLKDRPASGSPQLFLSKVEE 720  
QY 721 TFQICICQELVFRPITTVCOHNVCKCLDRSFRAQVFCPCRYDLGRSYAMOVNQPLQT 780  
Db 721 TFQICICQELVFRPITTVCOHNVCKCLDRSFRAQVFCPCRYDLGRSYAMOVNQPLQT 780

## RESULT 9

ABR82238  
ID ABR82238 standard; protein; 802 AA.  
AC ABR82238;  
XX  
XX  
DT 13-OCT-2003 (first entry)  
XX  
XX Human nucleic acid-associated protein (NAAP)-Id 4706628CD1.  
DE  
XX NAAAP; nucleic acid-associated protein; cardiant; cytostatic; transgenic;  
XX neuroprotective; gene therapy; human; cancer.  
KW  
XX Homo sapiens.  
XX  
XX W02003052048-A2.  
XX  
XX 26-JUN-2003.  
XX  
XX 02-MAY-2002; 2002WO-US014276.  
XX  
XX 04-MAY-2001; 2001US-0288598P.  
XX  
XX 17-MAY-2001; 2001US-0291776P.  
XX  
XX 18-MAY-2001; 2001US-0292172P.  
XX  
XX 25-MAY-2001; 2001US-0293564P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Yue H, Ding L, Baughn MR, Lal PG, Yue H, Hafalia AJA, Lee EA;  
XX Ison CH, Becha SD, Gururajan R, Emerling BM, Griffin JA, Tang YT;  
XX Lu DAM, Yao MG, Chawla NK, Ramkumar J, Gandhi AR, Lee SY;  
XX Richardson TW, Yang J, Elliott VS, Lu Y, Thangavelu K, He A;  
XX Azimzai Y, Raumann BE, Swarnakar A, Burford N;  
XX  
XX WPI; 2003-541640/51.  
XX  
XX N-FSDB; ACF35646.  
XX  
XX New human nucleic acid-associated proteins polypeptide, useful for  
XX preparing a composition for diagnosing or treating e.g., cardiovascular  
XX or neurological disorders.  
XX  
XX Claim 1; Page 168-169; 212pp; English.  
XX  
XX The invention relates to human nucleic acid-associated proteins (NAAAP)  
XX and encoding polynucleotides. The NAAAP polypeptides can be expressed by  
XX CC

CC standard recombinant methodology. The polypeptides are useful for  
CC preparing a composition for diagnosing or treating a disease or condition  
CC associated with decreased expression or overexpression of functional NAAAP  
CC e.g., cardiovascular or neurological disorders or cancer. The present  
CC sequence represents a human NAAAP polypeptide  
XX  
SQ Sequence 802 AA;

Query Match 53.7%; Score 2293.5; DB 6; Length 802;  
Best Local Similarity 53.3%; Pred. No. 7e-195;  
Matches 444; Conservative 126; Mismatches 192; Indels 71; Gaps 14;

QY 1 MWIQVTRDGRQTHVTDSLSRLTKVEELRRKIQELFHFVPEGLQRLFVRGQMEDGHTLPD 60  
Db 1 MWIQVTRDGRQTHVTDSLSRLTKVEELRRKIQELFHFVPEGLQRLFVRGQMEDGHTLPD 60  
QY 61 YEVRINDTTLQVRSQSL-VLPSTKEDSELSDTDSGCCIGQSESDKSTHGAAAEATDS 119  
Db 61 YDVGINDTTLQVRSQSL-VLPSTKEDSELSDTDSGCCIGQSESDKSTHGAAAEATDS 119  
QY 120 RPAEDMWDDELGLYKVEYVDARTDNTMGAFEAQVVRVTR-----KAFSRD----- 167  
Db 120 RPAEDMWDDELGLYKVEYVDARTDNTMGAFEAQVVRVTR-----KAFSRD----- 167  
QY 115 ARA-----RLDIPGFGIYKVELVDARDVGLGAFEAHHSVTRASDGOSRGKTPKNGSSC 171  
Db 115 ARA-----RLDIPGFGIYKVELVDARDVGLGAFEAHHSVTRASDGOSRGKTPKNGSSC 171  
QY 168 -----EPCSTSRP-----ALBEDVIYHVYDDYPENGVVQVNSRDVRA 206  
Db 168 -----EPCSTSRP-----ALBEDVIYHVYDDYPENGVVQVNSRDVRA 206  
QY 172 KRTNGNIKKHSKENTNKLDSPSTNSDCVAEDDVYHIQYDEYPSGTELMNVKDLRP 231  
Db 172 KRTNGNIKKHSKENTNKLDSPSTNSDCVAEDDVYHIQYDEYPSGTELMNVKDLRP 231  
QY 207 RARTILKQDLLEVQVVMNLYNPDNPKERGFWDALISLR-KRETRARELYANVLG--D 263  
Db 207 RARTILKQDLLEVQVVMNLYNPDNPKERGFWDALISLR-KRETRARELYANVLG--D 263  
QY 232 RARTILKQDLLEVQVVMNLYNPDNPKERGFWDALISLR-KRETRARELYANVLG--D 291  
Db 232 RARTILKQDLLEVQVVMNLYNPDNPKERGFWDALISLR-KRETRARELYANVLG--D 291  
QY 264 DSLNDCKRIIFVDEVFKIERPGEPSM--VDNPMRRKSGPCCKDKNVRLCRVCACHLC 321  
Db 264 DSLNDCKRIIFVDEVFKIERPGEPSM--VDNPMRRKSGPCCKDKNVRLCRVCACHLC 321  
QY 292 GTLNDCKLIISVDIFKIERFG-AHPLSFADGKFLRRNDPECDLGGDPEKKCHSCSRVC 350  
Db 292 GTLNDCKLIISVDIFKIERFG-AHPLSFADGKFLRRNDPECDLGGDPEKKCHSCSRVC 350  
QY 322 GGRQDPDKQLMCDCEMDFHIYCLDPLSSVPSDEWYCEPCRNDADEVLAGRLRESK 381  
Db 322 GGRQDPDKQLMCDCEMDFHIYCLDPLSSVPSDEWYCEPCRNDADEVLAGRLRESK 381  
QY 351 GGRHEPNQLLDECNVVHIYCLNPLPDKVPBEEYWCPSCKTDSSEVVKAGERLKXSK 410  
Db 351 GGRHEPNQLLDECNVVHIYCLNPLPDKVPBEEYWCPSCKTDSSEVVKAGERLKXSK 410  
QY 382 KNAKMASATSSORDMGKGMACVGRTECTIVPSNHYGPIPGIPVGTMMRFVQVSESGV 441  
Db 382 KNAKMASATSSORDMGKGMACVGRTECTIVPSNHYGPIPGIPVGTMMRFVQVSESGV 441  
QY 411 KKAQMPASSTESRRDWRGMACVGRTECTIVPSNHYGPIPGIPVGTMMRFVQVSESGV 470  
Db 411 KKAQMPASSTESRRDWRGMACVGRTECTIVPSNHYGPIPGIPVGTMMRFVQVSESGV 470  
QY 442 HRPHVAGIHGRSNDGSYSVLVAGGYEDVDHGNFTYTGSGGRDLGSKRTAEQSCDQKL 501  
Db 442 HRPHVAGIHGRSNDGSYSVLVAGGYEDVDHGNFTYTGSGGRDLGSKRTAEQSCDQKL 501  
QY 471 HRPHVAGIHGRSNDGSYSVLVAGGYEDVDHGNFTYTGSGGRDLGSKRTAEQSCDQKL 530  
Db 471 HRPHVAGIHGRSNDGSYSVLVAGGYEDVDHGNFTYTGSGGRDLGSKRTAEQSCDQKL 530  
QY 502 TTNTRALALNCAPINDQEGAEAKDWRSGKPVVRNVKGGKNSKYAPAEGRNYDGIYKV 561  
Db 502 TTNTRALALNCAPINDQEGAEAKDWRSGKPVVRNVKGGKNSKYAPAEGRNYDGIYKV 561  
QY 531 TTNTRALALNCAPINDQEGAEAKDWRSGKPVVRNVKGGKNSKYAPAEGRNYDGIYKV 590  
Db 531 TTNTRALALNCAPINDQEGAEAKDWRSGKPVVRNVKGGKNSKYAPAEGRNYDGIYKV 590  
QY 562 VKYWPEKGS-KFLVWRYLLRRDDDEPGPTWTEGKDRKKLGLTMQYPEGYLEALANRER 620  
Db 562 VKYWPEKGS-KFLVWRYLLRRDDDEPGPTWTEGKDRKKLGLTMQYPEGYLEALANRER 620  
QY 591 VKYWPEISSHGHFLVWRYLLRRDDDEPGPTWTEGKDRKKLGLTMQYPEGYLEALANRER 645  
Db 591 VKYWPEISSHGHFLVWRYLLRRDDDEPGPTWTEGKDRKKLGLTMQYPEGYLEALANRER 645  
QY 621 EKENSKEEBEQEGGFASPRTGKWKRSAGGSPSRAGSPRTSKTKVPEYSLTAQO 680  
Db 621 EKENSKEEBEQEGGFASPRTGKWKRSAGGSPSRAGSPRTSKTKVPEYSLTAQO 680  
QY 646 EGKPKGSKKQPSGTTKRP-----ISDDCPGASKVYKASDSAEATEAQLTPQQ 696  
Db 646 EGKPKGSKKQPSGTTKRP-----ISDDCPGASKVYKASDSAEATEAQLTPQQ 696  
QY 681 SSLIREDKSNKLNWELASLKDRPASGSPQLFLSKVEETFCICCOELVFRPITTVCO 740  
Db 681 SSLIREDKSNKLNWELASLKDRPASGSPQLFLSKVEETFCICCOELVFRPITTVCO 740  
QY 697 QHLIREDCQOKLWDELVSHLVEGPN-----FLKKLEQSFMCVCCOELVYQVPTTECF 749  
Db 697 QHLIREDCQOKLWDELVSHLVEGPN-----FLKKLEQSFMCVCCOELVYQVPTTECF 749  
QY 741 HNVCCKCLDRSFRAQVFCPCRYDLGRSYAMOVNQPLQTLVNLQFPYGNR 793  
Db 741 HNVCCKCLDRSFRAQVFCPCRYDLGRSYAMOVNQPLQTLVNLQFPYGNR 793

## RESULT 10

ABU69599  
ID ABU69599 standard; protein; 645 AA.  
XX



signal transduction; hyperproliferative disorder; diabetes mellitus; vitamin B12 malabsorption; neurological disorder; Huntington's chorea; Turner's syndrome; bacterial infection; cardiovascular disorder; infertility; psoriasis; haemolytic anaemia; anti-inflammatory; anti-HIV; cytostatic; hepatotropic; virucide; antirheumatic; antiarthritic; antiasthmatic; immunomodulator; antidiabetic; antiallergic; neuroprotective; immunosuppressive; vulnerable; antibacterial; antinfertility; antianaemic; antipsoriatic; cerebroprotective; cardiant; antiarteriosclerotic.

Homo sapiens.

WO200286076-A2.

31-OCT-2002.

19-APR-2002; 2002WO-US012636.

19-APR-2001; 2001US-0284962P.

26-APR-2001; 2001US-0286645P.

09-JAN-2002; 2002US-0346986P.

(BRIM ) BRISTOL-MYERS SQUIBB CO.

Carman J, Feder J, Nadler S;

WPI; 2003-0931119/08.

N-PSDB; ACA54638.

Novel NF-kappaB-associated polypeptides and polynucleotides useful for diagnosing, treating and preventing cancer, hepatic disorders, aberrant apoptosis, viral infections, autoimmune disorders, asthma and stroke.

Claim 4; Page 497-499; 608pp; English.

The present invention relates to the isolation of human nuclear factor-kappaB (NF-kappaB) associated polypeptides and polynucleotides. The NF-kappaB associated polypeptide and polynucleotide sequences are useful for preventing, treating or ameliorating various disorders including to disorders, inflammatory disorders, cancers, disorders relating to aberrant apoptosis, hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IGM syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic ectodermal dysplasia, immunodeficiency, al incontinentia pigmenti, viral infections (e.g. those caused by human immunodeficiency virus (HIV), human T-cell lymphotropic virus (HTLV), hepatitis B, hepatitis C, Epstein Barr virus (EBV), influenza), rheumatoid arthritis, inflammatory bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick syndrome, stroke, experimental allergic encephalomyelitis (EAE), autoimmune disorders, disorders related to hyper immune activity, disorders related to aberrant acute phase responses, hypercongenital conditions, birth defects, necrotic lesions, wounds, organ transplant rejection, disorders related to aberrant signal transduction, hyperproliferative disorders, diseases of the pancreas (e.g. diabetes mellitus, vitamin B12 malabsorption), neurological disorders (e.g. Huntington's chorea), Turner's syndrome, bacterial infections, cardiovascular disorders, infertility, psoriasis and haemolytic anaemia. The present sequence represents a human NF-kappaB associated polypeptide of the invention

Sequence 645 AA;

Query Match 43.8%; Score 1872; DB 6; Length 645;  
Best Local Similarity 55.8%; Pred. No. 2.1e-157;  
Matches 359; Conservative 97; Mismatches 137; Indels 50; Gaps 11;

Qy 1 MWIQRVWDGQTHVDSLSRLTKVEELRRKIQELFHVPEGLQRLFVRGKQMEGHTLFD 60  
Dy 1 MWIQRVTDGSKTCTIEDVSRKATIELRLVRWALFVRPEQRLFVRGKQLNGYTLFD 60  
Qy 61 YEVELNDTIQLVRQSL-VLPHSYKERDSELSLTDSCCLGQSESDKSTHGEAAETDS 119  
Dy 61 YDVGNDIQLVLRPDPDHLPGTGTQTEAK-----PCSNPPKVKKAPRVGPSNPST 114

Qy 120 RPAEDMWDETELGLYKXVYVDARDTDMGAWFEAOVVRVTR-----KAPSRD----- 167  
Dy 115 ARA---RLIDPGGIYKVNELVDARDVGLGAWFEAHHSVTRASDGGSRGKTPLKNGSSC 171  
Qy 168 -----EPCSTSRP---ALBEDVIYHVKYDDYPENGVMNSRDVRA 206  
Dy 172 KRTNGNIKHKS KENTNKLDSPSTNSDCVADEDDVIYHIQYDEYPSGTLEMMVKOLRP 231  
Qy 207 RARTIKQDLEVGQVVMNPDNPKERGFWDAEISR-KRETREARELYANVVLG--D 263  
Dy 232 RARTILKWNELNVGDVVMVNVNVEPCQGFWEFDEAITTLKTSRTKKELRVKIFLGGE 291  
Qy 264 DSLNDCELIIEVDEVEKIERPEGSEPM--VDMPMRKSGPSCKKCKODVNLRCVCACHLC 321  
Dy 292 GTLNDCKLIISVDEIFKIEREG-AHPJSFADGKFLRRNDPECDLCGGDEPKKCHSCSRVC 350  
Qy 322 GGRQDDPKQLMCDCEMFAHIYCLDPLSLSPSEDEWYCEPCRNDASEVVLAGERLRESK 381  
Dy 351 GGRHEPNMQLLDCENVAHYIYCLNPPLDKVPEEYWCPSCKTDSSEVVKAGERLKWSK 410  
Qy 382 KNAKMASATSSQDQWKGKMACVGRTECTIVPSNHYGPIPGIPVGTWPRFRVOVSEGV 441  
Dy 411 KKAKMPSASTESRRDNGRGMACVGRTECTIVPSNHYGPIPGIPVGTWPRFRVOVSEAGV 470  
Qy 442 HRPHVAGIHGRSNDGYSYSLVLAGYEDVDHGNPFYTGSGRDLNKGKRTAEQSCDQKL 501  
Dy 471 HRPHVGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGNLGNKRIKIGAFSAQTL 530  
Qy 502 TTNRALALNCFAPINDQGAEAKEKWSGKPRVVRNVKGGKNSKYAPAGENRYDGIYKV 561  
Dy 531 TNNRALALNCDAPLDDKIGAESRNWRGAKPVIRSFKGRKISKYAPEEGNRYDGIYKV 590  
Qy 562 VKYWEPEKGS-GLVWRYLLRRDDDEPGWTKGKDRIKKLG 603  
Dy 591 VKYWEPSSSHGLVWRYLLRRDDVEPAPWTSEGIERSRLCL 633

RESULT 12  
ABB76983  
ID ABB76983 standard; protein; 174 AA.  
XX ABB76983;  
XX AC ABB76983;  
DT 22-JUL-2002 (first entry)  
XX Human Inverted CCAAT box binding protein, ICBP90, fragment #3.  
DE Human Inverted CCAAT box binding protein; ICBP90; cytostatic;  
KW Human; inverted CCAAT box binding protein; ICBP90; cytostatic;  
KW cell proliferation control; inverted CCAAT box; cancer.  
XX Homo sapiens.  
XX WO200078949-A1.  
XX 28-DEC-2000.  
XX 22-JUN-2000; 2000WO-FR001747.  
XX 22-JUN-1999; 99FR-00007935.  
XX (ADER-) ADEREGEM ASSOC DEV RECH EN GENETIQ.  
XX Bronner C, Hopfner R, Mousli M, Jeltsch J, Lutz Y, Oudet P;  
XX WPI; 2001-091571/10.  
XX N-PSDB; ABL58023.  
XX Novel inverted CCAAT box binding protein, and related nucleic acids,  
XX antibodies and specific ligands, useful for treating and preventing  
XX cancer.  
XX Claim 2; Page 103; 115pp; French.

CC The present sequence is a protein fragment of human ICBP90 (inverted  
CC CCAAT box binding protein). The inverted CCAAT box is implicated in cell  
CC proliferation control. Several copies of the inverted CCAAT box are  
CC present in the promoter of the topoisomerase IIalpha gene, and also  
CC functions as a nuclear receptor. ICBP90 (ABP64013) and its coding  
CC sequence (ABL58020) are useful for treatment and/or prevention of cancer  
XX  
SQ Sequence 174 AA;

Query Match 21.7%; Score 926; DB 4; Length 174;  
Best Local Similarity 100.0%; Pred. No. 7.3e-74;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 620 REKENSREBEEOEGGFASPRTGKWKRSKSGGSPRAGSPRRTSKTKVPEPYSILTAQ 679  
Db 1 REKENSREBEEOEGGFASPRTGKWKRSKSGGSPRAGSPRRTSKTKVPEPYSILTAQ 60  
QY 680 QSSLIREDKSNKLMNEVLASLKDPRASGSPFOLFSLKVEETFCQICCOELVFRPITTV 739  
Db 61 QSSLIREDKSNKLMNEVLASLKDPRASGSPFOLFSLKVEETFCQICCOELVFRPITTV 120  
QY 740 QHNVCXCDLDRFRAQVFCPCACRYDLGRSYAMQVNOPLQTVLNQLPFGYNGR 793  
Db 121 QHNVCXCDLDRFRAQVFCPCACRYDLGRSYAMQVNOPLQTVLNQLPFGYNGR 174

RESULT 13  
ABP64013  
ID ABP64013 standard; protein; 198 AA.  
XX  
AC ABP64013;  
DT  
DT 04-NOV-2002 (first entry)  
XX  
DE Human ORF383.  
XX  
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
KW Anti-inflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
KW cancer; cardiovascular disease; allergy; autoimmune disease;  
KW wound healing; blood coagulation disorder; inflammatory disorder.  
XX  
OS Homo sapiens.  
XX  
XX US2002082206-A1.  
XX  
XX 27-JUN-2002.  
XX  
XX 30-MAY-2001; 2001US-00867550.  
XX  
XX 30-MAY-2000; 2000US-0208427P.  
XX  
XX (LEAC/) LEACH M D.  
XX (MEHR/) MEHRABAN F.  
XX (CONL/) CONLEY P B.  
XX (TOPP/) TOPPER J N.  
XX (LAWD/) LAW D.

Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
WPI; 2002-626554/67.  
N-PSDB; ABQ98576.  
XX  
XX New polypeptide designated ORFX are present in human atherogenic cells  
XX and are useful to prevent and treat ORFX-associated disorders including  
XX cancer, allergy, wound healing or autoimmune, cardiovascular or  
XX inflammatory disease.  
XX  
XX Claim 10; SEQ ID NO 766; 78pp; English.  
XX  
XX The present invention relates to novel human ORFX polypeptides and their  
XX coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99326). The sequences  
XX were discovered in human atherogenic cells, in particular in platelets

CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
CC many other tissues as well. Atherogenic cells are cells which have the  
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
CC nucleic acids are useful for treating or preventing a pathological  
CC condition associated with an ORFX-associated disorder, e.g. cancer,  
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
CC coagulation disorders or inflammatory disorders. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from the USPTO web site at  
CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
XX  
SQ Sequence 198 AA;

Query Match 19.3%; Score 823.5; DB 5; Length 198;  
Best Local Similarity 75.8%; Pred. No. 1.3e-64;  
Matches 150; Conservative 22; Mismatches 25; Indels 1; Gaps 1;  
QY 384 AKMASATSSORDWKGMACVGRTECTIVPSNHYGPIGIPVGTWRRFVQVSEGVHR 443  
Db 1 AKMPSASTESRRDWRGMACVGRTECTIVPSNHYGPIGIPVGTWRRFVQVSEGVHR 60  
QY 444 PHVAGIHGRSNDGSLVLGGVDDVDHGNFTYTGSGRDLGSGNKRATGSCDOKLTN 503  
Db 61 PHVGGIHGRSNDGSLVLGGVDDVDHGNFTYTGSGRDLGSGNKRATGSCDOKLTN 120  
QY 504 TNRALALNCFAPINDQEGAEAKDWRSGKFPVRVNRNVKGGKSKYAPAEGRNRYDGIYKVK 563  
Db 121 MNRALALNCDAPLDDKIGAESNRWAGKFPVRVIRSPKGRKISKYAPEEGNRYDGIYKVK 180  
QY 564 YWPEKGKS-GFLVWRYLL 580  
Db 181 YWPEISSHGHFLVWRYLL 198

RESULT 14  
AAU16348  
ID AAU16348 standard; protein; 133 AA.  
XX  
AC AAU16348;  
DT  
DT 07-NOV-2001 (first entry)  
XX  
DE Human novel secreted protein, Seq ID 1301.  
XX  
XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;  
KW antibacterial; virucide; fungicide; opthalmological; vulnerary;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
KW preservative; antiproliferative.  
XX  
XX Homo sapiens.  
XX  
XX WO200155322-A2.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US0001341.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 24-FEB-2000; 2000US-0184664P.  
XX 02-MAR-2000; 2000US-0186350P.  
XX 16-MAR-2000; 2000US-0189874P.  
XX 17-MAR-2000; 2000US-0190076P.  
XX 18-APR-2000; 2000US-0198123P.  
XX 19-MAY-2000; 2000US-0205515P.  
XX 07-JUN-2000; 2000US-0209467P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 18-AUG-2000; 2000US-0225759P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-488783/53.  
XX N-PSDB; AAS26335.  
XX New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.  
XX Claim 11; SEQ ID NO 1301; 980pp; English.  
XX The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic

immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence represents a novel secreted protein of the invention. Note: The sequence data for this patent did not form part of the printed

Query Match 15.7%; Score 672; DB 4; Length 133;  
 Best Local Similarity 99.2%; Pred. No. 2.2e-51;  
 Matches 126; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 667 KTKVPEYSLTAQSSLIREDKSNKLNVEVLASLKDPRASGSPFQLFSKVEETFCIC 726  
 Db 7 KTKVPEYSLTAQSSLIREDKSNKLNVEVLASLKDPRASGSPFQLFSKVEETFCIC 66  
 QY 727 QBELVFRPITTVCOHNVCKDCLDRSPFAOVFSCPACRYDLGRSYAMQVNPLOTVLNQLF 786  
 Db 67 QBELVFRPITTVCOHNVCKDCLDRSFXAQVFCPACRYDLGRSYAMQVNPLOTVLNQLF 126  
 QY 787 PGYNGNR 793  
 Db 127 PGYNGNR 133

RESULT 15  
 ABUS5417  
 ID ABUS5417 standard; protein; 133 AA.

AC ABUS5417;  
 DT 18-MAR-2003 (first entry)  
 XX Human novel polypeptide #504.

KW Human; neural disorder; immune system disorder; renal disorder;  
 KW muscular disorder; respiratory disease; reproductive disorder;  
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
 KW haemostatic; antiarteriosclerotic.

OS Homo sapiens.

XX US2002132753-A1.

XX 19-SEP-2002.

XX 17-JAN-2001; 2001US-00764864.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 28-JUN-2000; 2000US-0214886P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 08-SEP-2000; 2000US-0229513P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.  
 (RUBE/) RUBEN S M.  
 (BARA/) BARASH S C.

PI Rosen CA, Ruben SM, Barash SC;

XX WPI; 2003-147444/14.

XX N-PSDB; ABX73676.

XX New polypeptides and nucleic acids, useful in gene therapy for treating, inhibiting or preventing e.g. neural, immune system, muscular, respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or renal disorders.

XX Claim 11; SEQ ID NO 1301; 402pp; English.

XX The invention relates to human novel polypeptides and their associated polynucleotides. The polypeptides and polynucleotides are useful in gene therapy for treating, inhibiting or preventing neural disorders, immune system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis), muscular disorders, respiratory diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders, gastrointestinal disorders, pulmonary disorders, cardiovascular disorders (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute kidney failure and end-stage renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and leukaemia), inflammatory diseases (e.g. septic shock, bursitis and appendicitis), allergic reactions and conditions (e.g. asthma), blood related disorders (e.g. thrombosis, atherosclerosis and myocardial infarction) and cancerous diseases. Sequences ABUS4914-ABUS5699 and ABUS5748 represent human novel polypeptides of the invention

XX Sequence 133 AA;

XX SQ

Tue Nov 2 11:31:38 2004

Query Match 15.7%; Score 672; DB 6; Length 133;  
Best Local Similarity 99.2%; Pred. No. 2.2e-51;  
Matches 126; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 667 KKTKEPYSLTACQSSLIREDKSNKLNWVLSLKDRLPASGSPFQLFLSKVETFOCIC 726  
Db 7 KKTKEPYSLTACQSSLIREDKSNKLNWVLSLKDRLPASGSPFQLFLSKVETFOCIC 66  
Qy 727 CQELVFRPITTVCOHNVCKDCLDRSFRAQVFCPACRYDLGRSYAMQVNOPLQTVLNQLF 786  
Db 67 CQELVFRPITTVCOHNVCKDCLDRSFRAQVFCPACRYDLGRSYAMQVNOPLQTVLNQLF 126  
Qy 787 PGYNGR 793  
Db 127 PGYNGR 133

Search completed: November 1, 2004, 15:43:25  
Job time : 96 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:50:13 ; Search time 26 Seconds  
(without alignments)  
2022.699 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 793

Sequence: 1 MWIQVTRMDGRQTHVDSLS.....VNQPLQTVLNQLFRGNGNR 793

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 478139 seqs, 66318000 residues

Word size : 0

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.1	422	4	US-09-252-991A-20583
2	8	1.0	269	4	US-09-107-532A-4608
3	8	1.0	336	4	US-09-543-681A-5330
4	8	1.0	484	4	US-09-242-913B-17
5	8	1.0	501	4	US-09-538-092-707
6	8	1.0	665	4	US-09-252-991A-28319
7	8	1.0	952	4	US-09-328-352-5611
8	7	0.9	24	4	US-09-066-330-6
9	7	0.9	48	4	US-09-270-767-59066
10	7	0.9	65	2	US-08-867-087B-30
11	7	0.9	68	4	US-09-252-991A-18367
12	7	0.9	75	4	US-09-732-210-833
13	7	0.9	87	4	US-09-248-796A-17666
14	7	0.9	92	4	US-09-248-796A-24553
15	7	0.9	102	4	US-09-252-991A-27417
16	7	0.9	104	4	US-09-252-991A-31308
17	7	0.9	114	4	US-09-690-454-59
18	7	0.9	129	4	US-09-883-777-4
19	7	0.9	133	4	US-09-252-991A-31158
20	7	0.9	135	4	US-09-270-767-33963
21	7	0.9	135	4	US-09-270-767-49180
22	7	0.9	140	4	US-09-270-767-48570
23	7	0.9	141	4	US-09-252-991A-24775
24	7	0.9	147	4	US-09-732-210-578
25	7	0.9	149	4	US-09-270-767-36338
26	7	0.9	149	4	US-09-270-767-51555
27	7	0.9	166	4	US-09-543-681A-5820

```

28 7 0.9 170 4 US-08-468-996-9 Sequence 9, Appli
29 7 0.9 172 4 US-09-252-991A-29480 Sequence 29480, A
30 7 0.9 176 4 US-09-252-991A-24805 Sequence 24805, A
31 7 0.9 177 3 US-09-199-637A-247 Sequence 247, App
32 7 0.9 178 4 US-09-765-111A-39 Sequence 39, Appl
33 7 0.9 189 4 US-09-216-393B-130 Sequence 130, App
34 7 0.9 194 4 US-09-252-991A-21199 Sequence 21199, A
35 7 0.9 198 4 US-09-252-991A-28720 Sequence 28720, A
36 7 0.9 210 4 US-09-252-991A-28014 Sequence 28014, A
37 7 0.9 223 4 US-09-710-279-674 Sequence 674, App
38 7 0.9 227 4 US-09-107-532A-3843 Sequence 3843, Ap
39 7 0.9 233 4 US-09-248-796A-20388 Sequence 20388, A
40 7 0.9 236 3 US-08-134-001C-3558 Sequence 3558, Ap
41 7 0.9 244 4 US-09-724-623-82 Sequence 82, Appl
42 7 0.9 247 4 US-09-252-991A-32129 Sequence 32129, A
43 7 0.9 249 4 US-09-248-796A-15815 Sequence 15815, A
44 7 0.9 250 2 US-08-867-087B-13 Sequence 13, Appl
45 7 0.9 272 4 US-09-252-991A-17461 Sequence 17461, A

```

#### ALIGNMENTS

```

RESULT 1
US-09-252-991A-20583
; Sequence 20583, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20583
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-20583

```

```

Query Match 1.1%; Score 9; DB 4; Length 422;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 615 LANREREKE 623
Db 383 LANREREKE 391

```

```

RESULT 2
US-09-107-532A-4608
; Sequence 4608, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESS: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; CORRESPONDENCE ADDRESS: 7310
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC

```

```

; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/107,532A
;   FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/085,598
;   FILING DATE: 14 May 1998
;   APPLICATION NUMBER: 60/051571
;   FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
;   NAME: Ariniello, Pamela Deneke
;   REGISTRATION NUMBER: 40,489
;   REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (781)893-5007
;   TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4608:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 269 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
;   ORGANISM: Enterococcus faecium
;   FEATURE:
;     NAME/KEY: misc feature
;     LOCATION: (B) LOCATION 1...269
;   SEQUENCE DESCRIPTION: SEQ ID NO: 4608:
US-09-107-532A-4608

Query Match 1.0%; Score 8; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 599 KKLGLTMQ 606
Db 169 KKLGLTMQ 176

RESULT 3
US-09-543-681A-5330
; Sequence 5330, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;   APPLICANT: GARY BRETON
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
;   FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5330
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5330

Query Match 1.0%; Score 8; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 556 DGIYKVKV 563
Db 252 DGIYKVKV 259

RESULT 4
US-09-242-913B-17
; Sequence 17, Application US/09242913B
```

```

; Patent No. 6551811
; GENERAL INFORMATION:
;   APPLICANT: FONTAINE, THIERRY
;   APPLICANT: HARTLAND, ROBERT
;   APPLICANT: MOUYNA, ISABELLE
;   APPLICANT: LATGE, JEAN-PAUL
;   TITLE OF INVENTION: METHOD FOR SORTING ANTIFUNGAL MOLECULES ACTING ON THE
;   TITLE OF INVENTION: GLUCANOSYLTRANSFERASE ACTIVITY
;   FILE REFERENCE: 05986-0007
; CURRENT APPLICATION NUMBER: US/09/242,913B
; CURRENT FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: PCT/FR97/01540
; PRIOR FILING DATE: 1997-08-29
; PRIOR APPLICATION NUMBER: 60/024,910
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-242-913B-17

Query Match 1.0%; Score 8; DB 4; Length 484;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 ASATSSSQ 394
Db 448 ASATSSSQ 455

RESULT 5
US-09-538-092-707
; Sequence 707, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;   APPLICANT: Giot, Loic
;   APPLICANT: Mansfield, Traci A.
;   TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;   FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 707
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YOL070C
US-09-538-092-707

Query Match 1.0%; Score 8; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 VDSLSRLT 23
Db 467 VDSLSRLT 474

RESULT 6
US-09-252-991A-28319
; Sequence 28319, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
```

```
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28319
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28319

Query Match      1.0%; Score 8; DB 4; Length 665;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 372 LAGERLRE 379
Db 601 LAGERLRE 608

RESULT 7
US-09-328-352-5611
; Sequence 5611, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5611
; LENGTH: 952
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5611

Query Match      1.0%; Score 8; DB 4; Length 952;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 778 LQTVLNQL 785
Db 866 LQTVLNQL 873

RESULT 8
US-09-066-330-6
; Sequence 6, Application US/09066330A
; Patent No. 6511666
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: Bhogal, Peter S.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
; FILE REFERENCE: Reynolds
; CURRENT APPLICATION NUMBER: US/09/066,330A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: PN 6275
; EARLIER FILING DATE: 1995-10-30
; EARLIER APPLICATION NUMBER: PCT/AU96/00673
; EARLIER FILING DATE: 1996-10-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 24
```

```
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-09-066-330-6

Query Match      0.9%; Score 7; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 SSTHGEA 113
Db 9 SSTHGEA 15

RESULT 9
US-09-270-767-59066
; Sequence 59066, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59066
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-59066

Query Match      0.9%; Score 7; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 ASATSSS 393
Db 40 ASATSSS 46

RESULT 10
US-08-867-087B-30
; Sequence 30, Application US/08867087B
; Patent No. 5990386
; GENERAL INFORMATION:
; APPLICANT: An, Gynheung
; TITLE OF INVENTION: GENES CONTROLLING FLORAL DEVELOPMENT
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klarquist Sparkman Campbell Leigh &
; ADDRESSEE: Whinston, LLP
; STREET: One World Trade Center
; STREET: 121 S.W. Salmon Street
; STREET: Suite 1600
; CITY: Portland
; STATE: Oregon
; COUNTRY: United States of America
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk, 3-1/2 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,087B
; FILING DATE: June 2, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/323,449
; FILING DATE: October 14, 1994
```

APPLICATION NUMBER: U.S. 08/485,981  
FILING DATE: June 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Alan. E.  
REGISTRATION NUMBER: 35,123  
REFERENCE/DOCKET NUMBER: 4630-47071  
TELEPHONE: (503) 226-7391  
TELEFAX: (503) 228-9446  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 65 amino acid residues  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-867-087B-30

Query Match 0.9%; Score 7; DB 2; Length 65;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 VEELRRK 31  
Db 59 VEELRRK 65

RESULT 11  
US-09-252-991A-18367  
Sequence 18367, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-07-27  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 18367  
LENGTH: 68  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18367

Query Match 0.9%; Score 7; DB 4; Length 68;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 GSPRRTS 666  
Db 58 GSPRRTS 64

RESULT 12  
US-09-732-210-833  
Sequence 833, Application US/09732210  
Patent No. 6573361  
GENERAL INFORMATION:  
APPLICANT: Bunkers, Greg J.  
APPLICANT: Liang, Jihong  
APPLICANT: Mittanck, Cindy A.  
APPLICANT: Seale, Jeffrey W.  
APPLICANT: Wu, Yonnie S.  
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
FILE REFERENCE: 38-21(15036)B  
CURRENT APPLICATION NUMBER: US/09/732,210  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: US 60/169,513  
PRIOR FILING DATE: 1999-12-07  
PRIOR APPLICATION NUMBER: US 60/169,340

PRIOR FILING DATE: 1999-12-07  
SEQ ID NO 833  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Mycobacterium bovis  
US-09-732-210-833

Query Match 0.9%; Score 7; DB 4; Length 75;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 ERLRESK 381  
Db 20 ERLRESK 26

RESULT 13  
US-09-248-796A-17666  
Sequence 17666, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 17666  
LENGTH: 87  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-17666

Query Match 0.9%; Score 7; DB 4; Length 87;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 628 EEEEQQE 634  
Db 78 EEEEQQE 84

RESULT 14  
US-09-248-796A-24553  
Sequence 24553, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 24553  
LENGTH: 92  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-24553

Query Match 0.9%; Score 7; DB 4; Length 92;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      661  SPRTSK 667
Db      33  SPRTSK 39
|||||

RESULT 15
US-09-252-991A-27417
; Sequence 27417, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27417
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27417

Query Match      0.9%; Score 7; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      658  RAGSPRR 664
Db      14  RAGSPRR 20
|||||

Search completed: November 1, 2004, 15:56:13
Job time : 27 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:55:14 ; Search time 83 Seconds  
(without alignments)

3097.628 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 793

Sequence: 1 MWIQVTRMDGRQTHVTDSLS.....VNQPLQTVLNQLFPGVGNGR 793

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1370721 seqs, 324215800 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Length	ID	Description
1	591	74.5	793	US-10-123-568-2
2	591	74.5	793	Sequence 2, Appli
3	578	72.9	780	Sequence 27, Appl
4	136	17.2	136	Sequence 156, App
5	86	10.8	133	Sequence 3, Appli
6	19	2.4	198	Sequence 1301, Ap
7	19	2.4	645	Sequence 766, App
8	19	2.4	645	Sequence 113, App
9	19	2.4	802	Sequence 113, App
10	15	1.9	438	Sequence 7, Appli
11	15	1.9	750	Sequence 263042
12	15	1.9	774	Sequence 156872
13	14	1.8	617	Sequence 180766
				Sequence 522, App

14	1.8	617	15	US-10-374-780A-2314	Sequence 2314, Ap
15	1.3	110	9	US-09-764-864-848	Sequence 848, App
16	1.3	178	9	US-09-764-864-1303	Sequence 1303, Ap
17	1.1	92	14	US-10-106-698-7355	Sequence 7355, Ap
18	1.1	402	15	US-10-389-647-623	Sequence 623, App
19	1.1	1240	15	US-10-114-270-62	Sequence 62, Appl
20	1.1	1284	16	US-10-408-765A-1008	Sequence 1008, Ap
21	1.0	120	16	US-10-767-701-47576	Sequence 47576, A
22	1.0	166	11	US-09-864-408A-2368	Sequence 2368, Ap
23	1.0	262	16	US-10-437-963-135349	Sequence 135349, A
24	1.0	272	15	US-10-282-122A-52590	Sequence 52590, A
25	1.0	272	15	US-10-425-114-41709	Sequence 41709, A
26	1.0	297	15	US-10-424-599-171858	Sequence 171858, A
27	1.0	354	15	US-10-282-122A-51196	Sequence 51196, A
28	1.0	356	15	US-10-085-198-64	Sequence 64, Appl
29	1.0	356	15	US-10-210-172-148	Sequence 148, App
30	1.0	356	15	US-10-451-168-74	Sequence 74, Appl
31	1.0	356	16	US-10-471-115-9	Sequence 9, Appli
32	1.0	370	15	US-10-451-168-73	Sequence 73, Appl
33	1.0	386	16	US-10-437-963-202924	Sequence 202924, A
34	1.0	444	15	US-10-425-114-69164	Sequence 69164, A
35	1.0	484	14	US-10-347-278-17	Sequence 17, Appl
36	1.0	484	14	US-10-347-252-17	Sequence 17, Appl
37	1.0	613	16	US-10-767-701-46064	Sequence 46064, A
38	1.0	694	15	US-10-425-114-39797	Sequence 39797, A
39	1.0	709	15	US-10-424-599-174378	Sequence 174378, A
40	1.0	797	16	US-10-437-963-163096	Sequence 163096, A
41	1.0	946	15	US-10-282-122A-44976	Sequence 44976, A
42	0.9	13	14	US-10-014-340-701	Sequence 701, App
43	0.9	24	14	US-10-229-066-6	Sequence 6, Appli
44	0.9	32	14	US-10-174-410-275	Sequence 275, App
45	0.9	53	15	US-10-424-599-222236	Sequence 222236, A

#### ALIGNMENTS

##### RESULT 1

US-10-123-568-2  
; Sequence 2, Application US/10123568  
; Publication No. US20030194713A1  
; GENERAL INFORMATION:  
; APPLICANT: Hitoshi, Yasumichi  
; APPLICANT: Jenkins, Yonchu  
; APPLICANT: Rigel Pharmaceuticals, Inc.  
; TITLE OF INVENTION: NP95: Methods of Assaying for Cell Cycle Modulators  
; FILE REFERENCE: 021044-003400US  
; CURRENT APPLICATION NUMBER: US/10/123.568  
; CURRENT FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 793  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human NP95 nuclear zinc finger protein  
US-10-123-568-2

Query Match	74.5%	Score 591;	DB 14;	Length 793;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 791;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	MWIVRTMDGRQTHVTDSLSRLTKVIELRRKIOELFHVPEGLQFLFYRGKQMEDGHTLFD 60		
Db	1	MWIVRTMDGRQTHVTDSLSRLTKVIELRRKIOELFHVPEGLQFLFYRGKQMEDGHTLFD 60		
Qy	61	YEVELNDTIQLLVRSQVLVPHSTKERDSELSDDTSGCCLQSESDKSTHGEAAAEATDSR 120		
Db	61	YEVELNDTIQLLVRSQVLVPHSTKERDSELSDDTSGCCLQSESDKSTHGEAAAEATDSR 120		
Qy	121	PAEDMDWDETLGLYKNEYVDVARDTNMGAWFAQVVRVTRKAFPSRDEPCSSTSRPALEE 180		

Db	121	PADEDMWDETELGLYKNEVVDARDTNMGAFEAQVVRVTRKAPSRDEPCSSTSRPALEE	180
Qy	181	DVIYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPNPKERGFWD	240
Db	181	DVIYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPNPKERGFWD	240
Qy	241	AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERGEGSPMVDNPMRRKSGP	300
Db	241	AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERGEGSPMVDNPMRRKSGP	300
Qy	301	SCKHCKDDVNRLCRVCACHLGGQDPDKQLMCECDMAFHIYCLDPPLSSVSEDEWYC	360
Db	301	SCKHCKDDVNRLCRVCACHLGGQDPDKQLMCECDMAFHIYCLDPPLSSVSEDEWYC	360
Qy	361	PECNDASEVVLAGERLRESKNAKMASATSSSQORDWKGKMACVGRTKETIIVPSNHYGP	420
Db	361	PECNDASEVVLAGERLRESKNAKMASATSSSQORDWKGKMACVGRTKETIIVPSNHYGP	420
Qy	421	IPGIPVGTMMRFVRVQVSESGVHRPHVAGIHGRSNDGSGSLVLAGGYEDDVHGNFFTYTG	480
Db	421	IPGIPVGTMMRFVRVQVSESGVHRPHVAGIHGRSNDGSGSLVLAGGYEDDVHGNFFTYTG	480
Qy	481	SGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK	540
Db	481	SGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK	540
Qy	541	GGKNSKYAPAEGRNYDGIYKVYKWPKEKSGFLVWRYLRRDDDEPGPMTKEGKDRIKK	600
Db	541	GGKNSKYAPAEGRNYDGIYKVYKWPKEKSGFLVWRYLRRDDDEPGPMTKEGKDRIKK	600
Qy	601	LGLTMQYPEGYLEALANREREKSKREBEQEGGFASPTGKGKWKRSAGGSPSRAG	660
Db	601	LGLTMQYPEGYLEALANREREKSKREBEQEGGFASPTGKGKWKRSAGGSPSRAG	660
Qy	661	SPRRTSKTKTVEPYSLTAAQSSLREDKSNKLNNEVLASLKDRPASGPPQLFLSKVEE	720
Db	661	SPRRTSKTKTVEPYSLTAAQSSLREDKSNKLNNEVLASLKDRPASGPPQLFLSKVEE	720
Qy	721	TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCRYDLGRSYAMQVNPLOT	780
Db	721	TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCRYDLGRSYAMQVNPLOT	780
Qy	781	VLNQLFPGYGNGR 793	
Db	781	VLNQLFPGYGNGR 793	
RESULT 2			
US-10-188-832-27			
; Sequence 27, Application US/10188832			
; Publication No. US20040076955A1			
; GENERAL INFORMATION:			
; APPLICANT: Mack, David H.			
; APPLICANT: Aziz, Natasha			
; APPLICANT: Eos Biotechnology, Inc.			
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions			
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder			
; TITLE OF INVENTION: Cancer			
; FILE REFERENCE: 018501-002330US			
; CURRENT APPLICATION NUMBER: US/10/188,832			
; CURRENT FILING DATE: 2002-11-22			
; PRIOR APPLICATION NUMBER: US 60/302,814			
; PRIOR FILING DATE: 2001-07-03			
; PRIOR APPLICATION NUMBER: US 60/310,099			
; PRIOR FILING DATE: 2001-08-03			
; PRIOR APPLICATION NUMBER: US 60/343,705			
; PRIOR FILING DATE: 2001-11-08			
; PRIOR APPLICATION NUMBER: US 60/350,666			
; PRIOR FILING DATE: 2001-11-13			
; PRIOR APPLICATION NUMBER: US 60/372,246			
; PRIOR FILING DATE: 2002-04-12			
; NUMBER OF SEQ ID NOS: 207			
; SOFTWARE: PatentIn Ver. 2.1			

; SEQ ID NO 27			
; LENGTH: 793			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-188-832-27			
Query Match 74.5%; Score 591; DB 15; Length 793;			
Best Local Similarity 99.7%; Pred. No. 0;			
Matches 791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	MTQVTRMDGRQHTVDSLSRLTKVBELRRKIKIQLFHFVEPGLQRLFYRGKOMEDGHTLFD	60
Db	1	MTQVTRMDGRQHTVDSLSRLTKVBELRRKIKIQLFHFVEPGLQRLFYRGKOMEDGHTLFD	60
Qy	61	YEVRLNDTTLQVRSIPLPHSTKERDSELSDDTSGCCLGQSESDKSSSTHGEAAAEADSR	120
Db	61	YEVRLNDTTLQVRSIPLPHSTKERDSELSDDTSGCCLGQSESDKSSSTHGEAAAEADSR	120
Qy	121	PADEDMWDETELGLYKNEVVDARDTNMGAFEAQVVRVTRKAPSRDEPCSSTSRPALEE	180
Db	121	PADEDMWDETELGLYKNEVVDARDTNMGAFEAQVVRVTRKAPSRDEPCSSTSRPALEE	180
Qy	181	DVIYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPNPKERGFWD	240
Db	181	DVIYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNPNPKERGFWD	240
Qy	241	AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERGEGSPMVDNPMRRKSGP	300
Db	241	AEISRKRETRTARELYANVVLGDDSLNDCRIIFVDEVFKIERGEGSPMVDNPMRRKSGP	300
Qy	301	SCKHCKDDVNRLCRVCACHLGGQDPDKQLMCECDMAFHIYCLDPPLSSVSEDEWYC	360
Db	301	SCKHCKDDVNRLCRVCACHLGGQDPDKQLMCECDMAFHIYCLDPPLSSVSEDEWYC	360
Qy	361	PECNDASEVVLAGERLRESKNAKMASATSSSQORDWKGKMACVGRTKETIIVPSNHYGP	420
Db	361	PECNDASEVVLAGERLRESKNAKMASATSSSQORDWKGKMACVGRTKETIIVPSNHYGP	420
Qy	421	IPGIPVGTMMRFVRVQVSESGVHRPHVAGIHGRSNDGSGSLVLAGGYEDDVHGNFFTYTG	480
Db	421	IPGIPVGTMMRFVRVQVSESGVHRPHVAGIHGRSNDGSGSLVLAGGYEDDVHGNFFTYTG	480
Qy	481	SGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK	540
Db	481	SGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRNVK	540
Qy	541	GGKNSKYAPAEGRNYDGIYKVYKWPKEKSGFLVWRYLRRDDDEPGPMTKEGKDRIKK	600
Db	541	GGKNSKYAPAEGRNYDGIYKVYKWPKEKSGFLVWRYLRRDDDEPGPMTKEGKDRIKK	600
Qy	601	LGLTMQYPEGYLEALANREREKSKREBEQEGGFASPTGKGKWKRSAGGSPSRAG	660
Db	601	LGLTMQYPEGYLEALANREREKSKREBEQEGGFASPTGKGKWKRSAGGSPSRAG	660
Qy	661	SPRRTSKTKTVEPYSLTAAQSSLREDKSNKLNNEVLASLKDRPASGPPQLFLSKVEE	720
Db	661	SPRRTSKTKTVEPYSLTAAQSSLREDKSNKLNNEVLASLKDRPASGPPQLFLSKVEE	720
Qy	721	TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCRYDLGRSYAMQVNPLOT	780
Db	721	TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCRYDLGRSYAMQVNPLOT	780
Qy	781	VLNQLFPGYGNGR 793	
Db	781	VLNQLFPGYGNGR 793	
RESULT 3			
US-10-295-027-156			
; Sequence 156, Application US/10295027			
; Publication No. US2003023250A1			
; GENERAL INFORMATION:			
; APPLICANT: Afar, Daniel			



APPLICANT: Aziz, Natasha  
APPLICANT: Ginsberg, Wendy M.  
APPLICANT: Gish, Kurt C.  
APPLICANT: Glynn, Richard  
APPLICANT: Hevezi, Peter A.  
APPLICANT: Mack, David H.  
APPLICANT: Murray, Richard  
APPLICANT: Watson, Susan R.  
APPLICANT: Eos Biotechnology, Inc.  
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
Methods of Screening for Modulators of Cancer  
FILE REFERENCE: 018501-012500US  
CURRENT APPLICATION NUMBER: US/10/295,027  
CURRENT FILING DATE: 2002-11-13  
PRIOR APPLICATION NUMBER: US 09/663,733  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: US 60/350,666  
PRIOR FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 60/335,394  
PRIOR FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: US 60/332,464  
PRIOR FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: US 60/334,393  
PRIOR FILING DATE: 2001-11-29  
PRIOR APPLICATION NUMBER: US 60/340,376  
PRIOR FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: US 60/347,211  
PRIOR FILING DATE: 2002-01-08  
PRIOR APPLICATION NUMBER: US 60/347,349  
PRIOR FILING DATE: 2002-01-10  
PRIOR APPLICATION NUMBER: US 60/355,250  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: US 60/356,714  
PRIOR FILING DATE: 2002-02-13  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1386  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 156  
LENGTH: 780  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-295-027-156

Query Match 72.9%; Score 578; DB 14; Length 780;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 778; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MWIQRVTDGROTHVDSLSRLTKVEELRRKIQELPHVEPGLQRLFYRGKQMGDHTLFD 60  
DB 1 MWIQRVTDGROTHVDSLSRLTKVEELRRKIQELPHVEPGLQRLFYRGKQMGDHTLFD 60

QY 61 YEVRINDTTLQLVROSLVPLSHSTKERSLSLSDSGCCILGQSESOKSSTHGEAAETDSR 120  
DB 61 YEVRINDTTLQLVROSLVPLSHSTKERSLSLSDSGCCILGQSESOKSSTHGEAAETDSR 120

QY 121 PADEMDWDETELGLYKVEYVDARTNMGAWFEAQVVRVTRKAPSRDEPCSTSRPALEE 180  
DB 121 PADEMDWDETELGLYKVEYVDARTNMGAWFEAQVVRVTRKAPSRDEPCSTSRPALEE 180

QY 181 DVYHVKYDDYPENGVOVQNSRDVARARTIIKWQDLEVGVOVMLNYPNDNPKRGFWYD 240  
DB 181 DVYHVKYDDYPENGVOVQNSRDVARARTIIKWQDLEVGVOVMLNYPNDNPKRGFWYD 240

QY 241 AETSRKRETRTARELYANVVLGDSLNDRCRIIFVDEVFKIERPGEQSPMVDNPMRKGSP 300  
DB 241 AETSRKRETRTARELYANVVLGDSLNDRCRIIFVDEVFKIERPGEQSPMVDNPMRKGSP 300

QY 301 SKCHCKDDVNRICRVACACHLGGRQDPDKQLMDECDMAFHLYCLDPLSSVPSDEWYVC 360  
DB 301 SKCHCKDDVNRICRVACACHLGGRQDPDKQLMDECDMAFHLYCLDPLSSVPSDEWYVC 360

QY 361 PECDNDASEVVLAGERLRESKKAKMASATSSSQORDWKGMACVGRTECTIVPSNHYP 420

DB 361 PECDNDASEVVLAGERLRESKKAKMASATSSSQORDWKGMACVGRTECTIVPSNHYP 420  
QY 421 IPGIPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSVLVLAGYEDDDVHGNFTYTG 480  
DB 421 IPGIPVGTMMRFRVQVSESGVHRPHVAGIHGRSNDGYSVLVLAGYEDDDVHGNFTYTG 480  
QY 481 SGGRLDSNKTABOSCDOKLTNTNRALALNCFPAPINDQEGAEAKDWSEKGPVVRVNVK 540  
DB 481 SGGRLDSNKTABOSCDOKLTNTNRALALNCFPAPINDQEGAEAKDWSEKGPVVRVNVK 540  
QY 541 GGKNSKYAPAEGRNYDGIYKVVWPEKSGFLVWRYLLRRDDDEPGWTKEGKDRICK 600  
DB 541 GGKNSKYAPAEGRNYDGIYKVVWPEKSGFLVWRYLLRRDDDEPGWTKEGKDRICK 600  
QY 601 LGLTMQYPEGYLEALANRERKENSKEEBEQEGGFASPRTGKWKWKKSAGGSPSRAG 660  
DB 601 LGLTMQYPEGYLEALANRERKENSKEEBEQEGGFASPRTGKWKWKKSAGGSPSRAG 660  
QY 661 SPRTSKTKVPEYSLTAQOSSLIREDKSNKAKLWNEVLASLKDRPASGPFOLFUSKVEE 720  
DB 661 SPRTSKTKVPEYSLTAQOSSLIREDKSNKAKLWNEVLASLKDRPASGPFOLFUSKVEE 720  
QY 721 TFOCICCOELVFRPITTVCOHNVCKDCLDRSFRAGVFCPCACRYDLGRSYAMQVNPLOT 780  
DB 721 TFOCICCOELVFRPITTVCOHNVCKDCLDRSFRAGVFCPCACRYDLGRSYAMQVNPLOT 780

## RESULT 4

US-10-123-568-3  
Sequence 3, Application US/10123568  
Publication No. US20030194713A1  
GENERAL INFORMATION:  
APPLICANT: Hitoshi, Yasumichi  
APPLICANT: Rigel Pharmaceuticals, Inc.  
TITLE OF INVENTION: NP95: Methods of Assaying for Cell Cycle Modulators  
FILE REFERENCE: 021044-003400US  
CURRENT APPLICATION NUMBER: US/10/123,568  
CURRENT FILING DATE: 2002-04-15  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 136  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:G1-2635  
US-10-123-568-3

Query Match 17.2%; Score 136; DB 14; Length 136;  
Best Local Similarity 100.0%; Pred. No. 3.2e-125;  
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 493 AFQSCDQKLTNTNRALALNCFPAPINDQEGAEAKDWSEKGPVVRVNVKGGKNSKYAPAG 552  
DB 1 AFQSCDQKLTNTNRALALNCFPAPINDQEGAEAKDWSEKGPVVRVNVKGGKNSKYAPAG 552

QY 553 NRYDGIYKVVWPEKSGFLVWRYLLRRDDDEPGWTKEGKDRICKLGLTMQYPEGYL 612  
DB 61 NRYDGIYKVVWPEKSGFLVWRYLLRRDDDEPGWTKEGKDRICKLGLTMQYPEGYL 612

QY 613 EALANRERKENSKEE 628  
DB 121 EALANRERKENSKEE 136

## RESULT 5

US-09-764-864-1301  
Sequence 1301, Application US/09764864  
Patent No. US20020132753A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1301
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (93)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1301

Query Match          10.8%; Score 86; DB 9; Length 133;
Best Local Similarity 100.0%; Pred. No. 5.7e-76;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 567 KKTVEPYSLTAQSSLIREDKSNAKLWNEVLASIKORPASGSPQLFLSKVEETFCIC 726
Db 7 KKTVEPYSLTAQSSLIREDKSNAKLWNEVLASIKORPASGSPQLFLSKVEETFCIC 66

Qy 727 CQELVFRPITTCQHNVCCKCLDRSF 752
Db 67 CQELVFRPITTCQHNVCCKCLDRSF 92

RESULT 6
US-09-867-550-766
; Sequence 766, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: Thereby
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 766
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (57)
; OTHER INFORMATION: wherein Xaa may be any one of Arg or Cys or Gly or Ser
US-09-867-550-766

Query Match          2.4%; Score 19; DB 9; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 409 ECTIVPSNHYGPIPGIPVG 427
Db 26 ECTIVPSNHYGPIPGIPVG 44

RESULT 7
US-10-126-103-113
; Sequence 113, Application US/10126103
; Publication No. US20030224486A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
```

```
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB PATHWAY
; FILE REFERENCE: D0108.np
; CURRENT APPLICATION NUMBER: US/10/126,103
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/284,962
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/286,645
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/346,986
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-126-103-113

Query Match          2.4%; Score 19; DB 14; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 409 ECTIVPSNHYGPIPGIPVG 427
Db 438 ECTIVPSNHYGPIPGIPVG 456

RESULT 8
US-10-431-096-113
; Sequence 113, Application US/10431096
; Publication No. US2004008696A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE REFERENCE: D0108A CIP
; CURRENT APPLICATION NUMBER: US/10/431,096
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: US 60/284,962
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 10/126,103
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/286,645
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/346,986
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 113
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-431-096-113

Query Match          2.4%; Score 19; DB 15; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 409 ECTIVPSNHYGPIPGIPVG 427
Db 438 ECTIVPSNHYGPIPGIPVG 456

RESULT 9
US-10-476-924-7
; Sequence 7, Application US/10476924
; Publication No. US20040152093A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry; DING, Li;
; APPLICANT: BAUGHN, Marian R.; LAL, Preeti G.;
; APPLICANT: BAUGHN, Huibin; HAFALIA, April J.A.;
; APPLICANT: LEE, Ernestine A.; ISON, Craig H.;
; APPLICANT: BECHA, Shanya D.; GURURAJAN, Rajagopal;
```

APPLICANT: EMERLING, Brooke M.; GRIFFIN, Jennifer A.;  
APPLICANT: TANG, Y. Tom; LU, Dyung Aina M.;  
APPLICANT: YAO, Monique G.; CHAWLA, Narinder K.;  
APPLICANT: RAMKUMAR, Jayalaxmi; GANDHI, Ameena R.;  
APPLICANT: LEE, Soo Yeun; RICHARDSON, Thomas W.;  
APPLICANT: YANG, Junning; ELLIOTT, Vicki S.;  
APPLICANT: LU, Yan; THANGAVELU, Kavitha;  
APPLICANT: HE, Ann; AZIMZAI, Valda;  
APPLICANT: RAUMANN, Brigitte E.; SWARNAKAR, Anita;  
APPLICANT: BURFORD, Neil  
TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS  
FILE REFERENCE: PF-0960 USN  
CURRENT APPLICATION NUMBER: US/10/476,924  
CURRENT FILING DATE: 2003-11-04  
PRIOR APPLICATION NUMBER: PCT/US02/14276  
PRIOR FILING DATE: 2002-05-02  
PRIOR APPLICATION NUMBER: US 60/288,598  
PRIOR FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: US 60/291,776  
PRIOR FILING DATE: 2001-05-17  
PRIOR APPLICATION NUMBER: US 60/292,172  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/293,564  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PERL Program  
SEQ ID NO 7  
LENGTH: 802  
TYPE: PRT  
ORGANISM: Homo sapiens  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No: 4706628CDI  
US-10-476-924-7

Query Match 2.4%; Score 19; DB 15; Length 802;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 409 ECTIVPSNHYGPIPGIPVG 427  
DB 438 ECTIVPSNHYGPIPGIPVG 456  
|||||

RESULT 10  
US-10-424-599-263042  
Sequence 263042, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 263042  
LENGTH: 438  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_79548C.1.pap  
US-10-424-599-263042

Query Match 1.9%; Score 15; DB 15; Length 438;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 478 YTGSGGRDLSGNKRT 492  
|||||

Db 313 YTGSGGRDLSGNKRT 327

## RESULT 11

US-10-437-963-156872  
Sequence 156872, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 156872  
LENGTH: 750  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_56499C.1.pap  
US-10-437-963-156872

Query Match 1.9%; Score 15; DB 16; Length 750;  
Best Local Similarity 100.0%; Pred. No. 2.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 478 YTGSGGRDLSGNKRT 492  
DB 331 YTGSGGRDLSGNKRT 345  
|||||

RESULT 12  
US-10-437-963-180766  
Sequence 180766, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 180766  
LENGTH: 774  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_78105C.1.pap  
US-10-437-963-180766

Query Match 1.9%; Score 15; DB 16; Length 774;  
Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 478 YTGSGGRDLSGNKRT 492  
DB 321 YTGSGGRDLSGNKRT 335  
|||||

```

; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2314
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G373
US-10-374-780A-2314

Query Match          1.8%; Score 14; DB 15; Length 617;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 478 YTGSGGRDLGNKR 491
Db 317 YTGSGGRDLGNKR 330

RESULT 15
US-09-764-864-848
; Sequence 848, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 848
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-848

Query Match          1.3%; Score 10; DB 9; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 754 AQVFSCPACR 763
Db 71 AQVFSCPACR 80

Search completed: November 1, 2004, 16:01:32
Job time : 84 secs

; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 522
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-522

Query Match          1.8%; Score 14; DB 14; Length 617;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 478 YTGSGGRDLGNKR 491
Db 317 YTGSGGRDLGNKR 330

RESULT 14
US-10-374-780A-2314
; Sequence 2314, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:46:28 ; Search time 104 Seconds

(without alignments)  
4387.232 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 793

Sequence: 1 MWIQVTRMDGRQTHVTDSL.....VNQPLQTLNQLFPGVGNGR 793

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	793	2	Q9P115
2	591	74.5	793	2	Q9T88
3	189	23.8	189	2	Q9H6S6
4	139	17.5	139	2	Q8J022
5	48	6.1	48	2	Q9P1U7
6	40	5.0	597	2	Q6IP39
7	40	5.0	597	2	AAH72079
8	32	4.0	782	2	Q9Z1H6
9	32	4.0	782	2	Q8VDF2
10	32	4.0	829	2	Q7TPK1
11	29	3.7	775	2	Q6PEI0
12	29	3.7	775	2	AAH58055
13	29	3.7	776	2	Q8CDP6
14	28	3.5	299	2	Q8C6F1
15	28	3.5	516	2	Q8BJP6
16	28	3.5	803	2	Q7TM13
17	28	3.5	803	2	Q8K1I5
18	28	3.5	803	2	AAH60241
19	20	2.5	474	2	Q8VIA1
20	19	2.4	503	2	Q8TAG7
21	19	2.4	802	2	Q96PU4
22	15	1.9	610	2	Q9FW25
23	15	1.9	641	2	Q9FVS3
24	15	1.9	645	2	Q8VYZ0
25	15	1.9	765	2	Q9VY20
26	15	1.9	789	2	Q7XW58
27	15	1.9	789	2	Q75M36
28	14	1.8	789	2	AAH88821
29	14	1.8	598	2	Q9C8E0
30	14	1.8	615	2	Q6NQ90
31	14	1.8	617	2	AAQ65191
					Q9FKA7
					Q9fka7 arabidopsis

32	14	1.8	617	2	AAQ65196
33	14	1.8	622	2	Q9C8E1
34	10	1.3	1858	2	Q8J0W7
35	10	1.3	1859	2	Q8J0Z1
36	10	1.3	1862	2	Q8J111
37	10	1.3	1863	2	Q8J0Y1
38	9	1.1	195	2	Q8BY30
39	9	1.1	216	2	Q8BG56
40	9	1.1	273	2	Q6TP33
41	9	1.1	273	2	AAZ27550
42	9	1.1	288	2	Q9BZ45
43	9	1.1	320	2	Q0S914
44	9	1.1	402	2	Q9HWG9
45	9	1.1	584	2	Q6JWV2

#### ALIGNMENTS

RESULT 1  
Q9P115  
ID Q9P115 PRELIMINARY; PRT; 793 AA.  
AC Q9P115;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Transcription factor ICBP90.  
GN Name=ICBP90;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20110783; PubMed=10646863;  
RA Hopfner R., Mousli M., Jeltsch J.M., Voulgaris A., Lutz Y., Marin C.,  
RA Bellocq J.P., Oudet P., Bronner C.;  
RT "ICBP90, a novel human CCAAT binding protein, involved in the  
RT regulation of topoisomerase I $\alpha$  expression.";  
RL Cancer Res. 60:121-128(2000).  
DR HSSP; Q9UIG0; 1F62  
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; TAS.  
DR GO; GO:0003700; F:transcription factor activity; TAS.  
DR GO; GO:0008283; P:cell proliferation; TAS.  
DR GO; GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.  
DR InterPro; IPR011011; FYVE\_PHD\_Znf.  
DR InterPro; IPR003105; Gqa.  
DR InterPro; IPR000626; Ubiquitin.  
DR InterPro; IPR001965; Znf\_PHD.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00628; PHD; 1.  
DR Pfam; PF00240; ubiquitin; 1.  
DR Pfam; PF02182; YDG\_SRA; 1.  
DR PRINTS; PR00348; UBIQUITIN.  
DR SMART; SM00249; PHD; 1.  
DR SMART; SM00184; RING; 2.  
DR SMART; SM00466; SRA; 1.  
DR SMART; SM00213; UBO; 1.  
DR PROSITE; PS00053; UBIQUITIN\_2; 1.  
DR PROSITE; PS01359; ZF\_PHD\_1; UNKNOWN\_1.  
DR PROSITE; PS00016; ZF\_PHD\_2; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; UNKNOWN\_1.  
DR PROSITE; PS00089; ZF\_RING\_2; 2.  
SQ SEQUENCE 793 AA; 59815 MW; D9B4161E892BB014 CRC64;

Query Match 100.0%; Score 793; DB 2; Length 793;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWIQVTRMDGRQTHVTDSLRLTKVEELRRKIQELFVHVEPGLQRLFYRGKMGEDGHTLFD 60

Db 1 MWIQVTRMDGRQTHVTDSLRLTKVEELRRKIQELFVHVEPGLQRLFYRGKMGEDGHTLFD 60



```

Db 601 LGITWQYPEGYLEALANREREKENSKEBEEQEGGFASPTGKGKWKXKSAGGSPSRAG 660
QY 661 SPRTSKTKTKVPYSYSLTAQSSLIREDKSNKLNWNEVLASLKDRPASGSPFQLFLSKVEE 720
Db 661 SPRTSKTKTKVPYSYSLTAQSSLIREDKSNKLNWNEVLASLKDRPASGSPFQLFLSKVEE 720
QY 721 TFCQCCQELVFRPITTVCOHNVCCKLDRSFRAQVFCPACRYDLGRSYAMQVNPLOT 780
Db 721 TFCQCCQELVFRPITTVCOHNVCCKLDRSFRAQVFCPACRYDLGRSYAMQVNPLOT 780
QY 781 VLNLQFPFGYNGNR 793
Db 781 VLNLQFPFGYNGNR 793

RESULT 3
QH6S6 PRELIMINARY; PRT; 189 AA.
AC Q9H6S6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ12925.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK025578; BAB15177.1; -
DR GO; GO:000151; Cubiquitin ligase complex; IEA.
DR GO; GO:0004842; Fubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; Zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN 1.
DR PROSITE; PS0089; ZF_RING_2; 1.
SQ SEQUENCE 189 AA; 23351 MW; 88C999C7029185AE CRC64;

Query Match 23.8%; Score 189; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.7e-186;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 MOYPEGYLEALANREREKENSKEBEEQEGGFASPTGKGKWKXKSAGGSPSRAG 664
Db 1 MOYPEGYLEALANREREKENSKEBEEQEGGFASPTGKGKWKXKSAGGSPSRAG 60
QY 665 TSXKTKVEPYSLTAQSSLIREDKSNKLNWNEVLASLKDRPASGSPFQLFLSKVEETFC 724
Db 61 TSXKTKVEPYSLTAQSSLIREDKSNKLNWNEVLASLKDRPASGSPFQLFLSKVEETFC 120
QY 725 ICCQELVFRPITTVCOHNVCCKLDRSFRAQVFCPACRYDLGRSYAMQVNPLOTVLNQ 784
Db 121 ICCQELVFRPITTVCOHNVCCKLDRSFRAQVFCPACRYDLGRSYAMQVNPLOTVLNQ 180
QY 785 LFPFGYNGNR 793
Db 181 LFPFGYNGNR 189

RESULT 4
Q8U022 PRELIMINARY; PRT; 139 AA.
AC Q8U022;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hn9p5 (Fragment).

```

```

OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Muto M., Kubo K., Kanari Y., Utsuno M., Matsuda Y., Tatsumi K.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB075601; BAC20576.1; -
FT NON_TER 1
FT NON_TER 139
SQ SEQUENCE 139 AA; 15653 MW; 07E795316304FBBD CRC64;

Query Match 17.5%; Score 139; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 8.2e-135;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 607 YPEGYLEALANREREKENSKEBEEQEGGFASPTGKGKWKXKSAGGSPSRAGS 666
Db 1 YPEGYLEALANREREKENSKEBEEQEGGFASPTGKGKWKXKSAGGSPSRAGS 60
QY 667 KTKVPEPYSLTAQSSLIREDKSNKLNWNEVLASLKDRPASGSPFQLFLSKVEETFCIC 726
Db 61 KTKVPEPYSLTAQSSLIREDKSNKLNWNEVLASLKDRPASGSPFQLFLSKVEETFCIC 120
QY 727 CQELVFRPITTVCOHNVCCK 745
Db 121 CQELVFRPITTVCOHNVCCK 139

RESULT 5
Q9PLU7 PRELIMINARY; PRT; 48 AA.
AC Q9PLU7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ICBP90 amino acids 746..793 (Fragment).
GS Name=ICBP90;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.B., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Attix C., Andreise T., Frankheim M., Amico-Keller G., Coefficient J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
RA Arellano A., Brower A., Sanders C., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC053467; AAF64067.1; -
FT NON_TER 1
SQ SEQUENCE 48 AA; 5457 MW; 7CCBCF47A9A32A4B CRC64;

Query Match 6.1%; Score 48; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 5.9e-41;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 746 DCLDRSFRAQVFCPACRYDLGRSYAMQVNPLOTVLNQLPFGYNGNR 793
Db 1 DCLDRSFRAQVFCPACRYDLGRSYAMQVNPLOTVLNQLPFGYNGNR 48

RESULT 6
Q6IP39 PRELIMINARY; PRT; 597 AA.
AC Q6IP39;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE LOC432234 protein (Fragment).
GN Name=LOC432234;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Udwin T.B., Toshiyuki S., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC072079; AAH72079.1; -.
DR InterPro; IPR011011; FYVE_PHD_ZnF.
DR InterPro; IPR003105; G9a.
DR InterPro; IPR000626; Ubiquitin.
DR InterPro; IPR001965; ZnF_PHD.
DR Pfam; PF00628; PHD; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR Pfam; PF02182; YDG_SRA; 1.
DR SMART; SM00249; PHD; 1.
DR SMART; SM00466; SRA; 1.
DR SMART; SM00213; UBO; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
DR PROSITE; PS01359; ZF_PHD 1; UNKNOWN_1.
DR PROSITE; PS50016; ZF_PHD 2; 1.
FT NON TER 597
SQ SEQUENCE 597 AA; 67054 MW; 273BEC791D9FA86E CRC64;

Query Match 5.0%; Score 40; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 9.8e-32;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 432 FRVQVSEGVHRPHVAGIHGRSNDGSYSVLVAGGYEDD 471
|
DB 430 FRVQVSEGVHRPHVAGIHGRSNDGSYSVLVAGGYEDD 469

RESULT 7
Query Match 5.0%; Score 40; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 9.8e-32;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 432 FRVQVSEGVHRPHVAGIHGRSNDGSYSVLVAGGYEDD 471
|
DB 430 FRVQVSEGVHRPHVAGIHGRSNDGSYSVLVAGGYEDD 469

AAH72079
ID AAH72079 PRELIMINARY; PRT; 597 AA.
AC AAH72079;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Udwin T.B., Toshiyuki S., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC072079; AAH72079.1; -.
DR InterPro; IPR003105; G9a.
DR InterPro; IPR000626; Ubiquitin.
DR InterPro; IPR001965; ZnF_PHD.
DR Pfam; PF00628; PHD; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR Pfam; PF02182; YDG_SRA; 1.
DR SMART; SM00249; PHD; 1.
DR SMART; SM00466; SRA; 1.
DR SMART; SM00213; UBO; 1.
DR PROSITE; PS50053; UBIQUITIN 2; 1.
DR PROSITE; PS01359; ZF_PHD 1; UNKNOWN_1.
DR PROSITE; PS50016; ZF_PHD 2; 1.
FT NON TER 597
SQ SEQUENCE 597 AA; 67054 MW; 273BEC791D9FA86E CRC64;

Query Match 5.0%; Score 40; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 9.8e-32;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 432 FRVQVSEGVHRPHVAGIHGRSNDGSYSVLVAGGYEDD 471
|
DB 430 FRVQVSEGVHRPHVAGIHGRSNDGSYSVLVAGGYEDD 469

RESULT 8
Q9Z1H6
ID Q9Z1H6 PRELIMINARY; PRT; 782 AA.
AC Q9Z1H6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Nuclear protein np95 (Nuclear zinc finger protein Np95).
GN Name=Uhrf1; Synonyms=Np95;

```





ID Q7TPK1 PRELIMINARY; PRT; 829 AA.  
 AC Q7TPK1;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE AC2-121.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Xu C.S., Li W.Q., Li Y.C., Wang G.P., Chai L.Q., Yuan J.Y., Yang K.J.,  
 RA Yan H.W., Chang C.F., Zhao L.F., Ma H., Wang L., Wang S.F., Han H.P.,  
 RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY321334; AAP86266.1; -  
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR003105; G9a  
 DR InterPro; IPR000586; Lipocln\_cytFAPB.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR InterPro; IPR001841; Znf\_PHD.  
 DR Pfam; PF00628; PHD; 1.  
 DR Pfam; PF00240; ubiquitin; 1.  
 DR Pfam; PF02182; YDC\_SRA; 1.  
 DR SMART; SM00249; PHD; 1.  
 DR SMART; SM00184; RING; 2.  
 DR SMART; SM00466; SRA; 1.  
 DR SMART; SM00213; UBQ; 1.  
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN\_1.  
 DR PROSITE; PS00053; UBIQUITIN\_2; 1.  
 DR PROSITE; PS01359; ZF\_PHD 1; UNKNOWN\_1.  
 DR PROSITE; PS00016; ZF\_PHD 2; 1.  
 DR PROSITE; PS00518; ZF\_RING 1; UNKNOWN\_1.  
 DR PROSITE; PS00089; ZF\_RING 2; 1.  
 SQ SEQUENCE 829 AA; 53222 MW; E6B8327F33FE74BE CRC64;  
 Query Match 4.0%; Score 32; DB 2; Length 829;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-23;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 425 PVGTWRFVQVSGVHRPHVAGIHGRSNDG 456  
 Db 477 PVGTWRFVQVSGVHRPHVAGIHGRSNDG 508  
 RESULT 11  
 ID Q6PEIO PRELIMINARY; PRT; 775 AA.  
 AC Q6PEIO;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Zgc:63539.  
 GN Name=zgc:63539;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=AB; TISSUE=Whole body;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Bouffard G.G.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB; TISSUE=Whole body;  
 RA Strausberg R.;  
 CC Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 DR EMBL; BC058055; AAH58055.1; -  
 DR InterPro; IPR011011; FYVE\_PHD\_Znf.  
 DR InterPro; IPR003105; G9a.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR InterPro; IPR001965; Znf\_PHD.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00628; PHD; 1.  
 DR Pfam; PF00240; ubiquitin; 1.  
 DR Pfam; PF02182; YDC\_SRA; 1.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR PRINTS; PR00348; UBIQUITIN.  
 DR SMART; SM00249; PHD; 1.  
 DR SMART; SM00184; RING; 2.  
 DR SMART; SM00466; SRA; 1.  
 DR SMART; SM00213; UBQ; 1.  
 DR PROSITE; PS00053; UBIQUITIN\_2; 1.  
 DR PROSITE; PS01359; ZF\_PHD 1; UNKNOWN\_1.  
 DR PROSITE; PS00016; ZF\_PHD 2; 1.  
 DR PROSITE; PS00518; ZF\_RING 1; UNKNOWN\_1.  
 DR PROSITE; PS00089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 SQ SEQUENCE 775 AA; 87090 MW; 96D60695EDA2468D CRC64;  
 Query Match 3.7%; Score 29; DB 2; Length 775;  
 Best Local Similarity 100.0%; Pred. No. 3e-20;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 552 GNRVDGIYKVKYWPCKSGFLVWRYLL 580  
 Db 549 GNRVDGIYKVKYWPCKSGFLVWRYLL 577  
 RESULT 12  
 AAH58055  
 ID AAH58055 PRELIMINARY; PRT; 775 AA.  
 AC AAH58055;  
 DT 20-MAY-2004 (TrEMBLrel. 27, Created)  
 DT 20-MAY-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Zgc:63539.  
 GN Zgc:63539.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB; TISSUE=Whole body;  
 RX MEDLINE=22388257; PubMed=12477932;



DR GO: GO:0000151; C:ubiquitin ligase complex; IEA.  
 DR GO: GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO: GO:0008270; F:zinc ion binding; IEA.  
 DR GO: GO:0016567; P:protein ubiquitination; IEA.  
 DR InterPro: IPR003105; Gnf.  
 DR InterPro: IPR001841; Znf\_ring.  
 DR Pfam: PF02182; YDG\_SRA; 1.  
 DR SMART: SM00184; RING; 1.  
 DR PROSITE: PS00518; ZF\_RING\_1; UNKNOWN\_1.  
 DR PROSITE: PS50089; ZF\_RING\_2; 1.  
 KW Nuclear protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 299 AA; 33647 MW; 512C0893DBAF048 CRC64;  
 Query Match 3.5%; Score 28; DB 2; Length 299;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-19;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 736 TTVCQHNVCCKDLRSFRAQVFSCPACR 763  
 ||||||||||||||||||||||||||||  
 Db 242 TTVCQHNVCCKDLRSFRAQVFSCPACR 269  
 RESULT 15  
 Q8BJP6 PRELIMINARY; PRT; 516 AA.  
 AC Q8BJP6;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length  
 DE enriched library, clone:B430210011 product:hypothetical Ubiquitin  
 DE domain containing protein, full insert sequence.  
 GN Name-Uhrf2;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Adipose;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RX Carninci P.; Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Adipose;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RX RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Adipose;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Adipose;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P.; Shibata Y.; Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Adipose;  
 RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kusunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Adipose;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK080925; BAC38081.1; -  
 DR HSSP: O9UIG0; 1F62.  
 DR MGD; MGI:1923718; Uhrf2.  
 DR GO: GO:0003677; F:DNA binding; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR011011; FYVE\_PHD\_Znf.  
 DR InterPro: IPR003105; G9A.  
 DR InterPro: IPR000626; Ubiquitin.  
 DR InterPro: IPR001965; Znf\_PHD.  
 DR Pfam: PF00628; PHD; 1.  
 DR Pfam: PF00240; ubiquitin; 1.  
 DR Pfam: PF02182; YDG\_SRA; 1.  
 DR SMART: SM00249; PHD; 1.  
 DR SMART: SM00213; UBO; 1.  
 DR PROSITE: PS50053; UBIQUITIN\_2; 1.  
 DR PROSITE: PS01359; ZF\_PHD\_1; UNKNOWN\_1.  
 DR PROSITE: PS50016; ZF\_PHD\_2; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 516 AA; 57761 MW; 3D6862E9D1E1DEDD CRC64;  
 Query Match 3.5%; Score 28; DB 2; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-19;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 400 GMACVGRTECTIVPSNHYGPIGPVG 427  
 ||||||||||||||||||||||||  
 Db 430 GMACVGRTECTIVPSNHYGPIGPVG 457  
 Search completed: November 1, 2004, 15:55:07  
 Job time : 106 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:49:58 ; Search time 28 Seconds  
(without alignments)

2724.996 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 793

Sequence: 1 MWIQVTRWDGRTHTVDSLS.....VNQPLQTVLNQLFFGYGNR 793

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 79:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	1.9	641	2 E96612	probable transcrip
2	14	1.8	598	2 A96685	probable RING zinc
3	14	1.8	622	2 H96684	probable RING zinc
4	9	1.1	402	2 E83119	probable PAD-depen
5	8	1.0	249	2 A82796	hypothetical prote
6	8	1.0	265	2 H97324	uncharacterized pr
7	8	1.0	294	2 S23561	HALL protein - yea
8	8	1.0	378	2 A82406	cytochrome d ubiq
9	8	1.0	394	2 T26968	hypothetical prote
10	8	1.0	484	2 S66713	hypothetical prote
11	8	1.0	501	2 S66763	hypothetical prote
12	8	1.0	1130	2 T21134	hypothetical prote
13	8	1.0	1587	2 G86467	hypothetical prote
14	7	0.9	64	1 S03778	uvib protein - clo
15	7	0.9	74	2 T42941	hypothetical prote
16	7	0.9	77	2 H70642	probable ribosomal
17	7	0.9	80	2 T45372	ribosomal protein
18	7	0.9	81	2 A87141	50S ribosomal prot
19	7	0.9	134	2 E91070	DNA-binding protei
20	7	0.9	134	2 E85914	DNA-binding protei
21	7	0.9	134	2 JH0774	DNA-binding protei
22	7	0.9	139	2 C90300	hypothetical prote
23	7	0.9	142	2 C72430	hypothetical prote
24	7	0.9	145	2 T17931	IgA Fc receptor-li
25	7	0.9	147	2 T35563	ribosomal protein
26	7	0.9	147	2 T26225	hypothetical prote
27	7	0.9	150	1 D71092	hypothetical prote
28	7	0.9	154	2 T28434	transcription regu
29	7	0.9	155	2 A60945	actin homolog FAI,

#### ALIGNMENTS

##### RESULT 1

E96612

probable transcription factor F12K22.14 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: E96612

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.; A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96612

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-641 <STO>

A:Cross-references: UNIPROT:Q9FVS3; GB:AE005173; NID:g11079528; PIDN:AGC29238.1; GSPDB:GN

C:Genetics:

A:Gene: F12K22.14

A:Map position: 1

C:Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

Query Match 1.9%; Score 15; DB 2; Length 641;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 478 YTGSGGRDLGNKKT 492

Db 332 YTGSGGRDLGNKKT 346

##### RESULT 2

A96685

probable RING zinc finger protein F15E12.5 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: A96685

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.; A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: A96685  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-598 <STO>  
A;Cross-references: UNIPROT:Q9C8E0; GB:AE005173; NID:gl1038479; PIDN:AAG27758.1; GSPDB:G  
C;Genetics:  
A;Gene: F15E12.5  
A;Map position: 1  
C;Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

Query Match 1.8%; Score 14; DB 2; Length 598;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 478 YTGSGGRDLGNKR 491  
|||||  
Db 292 YTGSGGRDLGNKR 305

RESULT 3  
H96684  
probable RING zinc finger protein F15E12.8 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: H96684  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: H96684  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-622 <STO>  
A;Cross-references: UNIPROT:Q9C8E1; GB:AE005173; NID:gl1038468; PIDN:AAG27747.1; GSPDB:G  
C;Genetics:  
A;Gene: F15E12.8  
A;Map position: 1  
C;Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

Query Match 1.8%; Score 14; DB 2; Length 622;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 478 YTGSGGRDLGNKR 491  
|||||  
Db 317 YTGSGGRDLGNKR 330

RESULT 4  
E83119  
probable FAD-dependent monooxygenase PA4217 [imported] - Pseudomonas aeruginosa (strain  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: E83119  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: E83119  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-402 <STO>  
A;Cross-references: UNIPROT:Q9HWG9; GB:AE004838; GB:AE004091; NID:g9950422; PIDN:AAG0760

A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA4217

Query Match 1.1%; Score 9; DB 2; Length 402;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 615 LANREREKE 623  
|||||  
Db 363 LANREREKE 371

RESULT 5  
A82796  
hypothetical protein XF0518 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: A82796  
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717; PMID:10910347  
A;Note: for a complete list of authors see reference number A59328 below  
A;Accession: A82796  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-249 <SIM>  
A;Cross-references: UNIPROT:Q9PFY8; GB:AE003900; GB:AE003849; NID:g9105366; PIDN:AAF8332f  
A;Experimental source: strain 9a5c  
R;Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carrer, H.  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to Genbank, June 2000  
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franco, S.C.; Franco, M.C.; Frohme  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.  
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A;Reference number: A59328  
A;Contents: annotation  
C;Genetics:  
A;Gene: XF0518

Query Match 1.0%; Score 8; DB 2; Length 249;  
Best Local Similarity 100.0%; Pred. No. 9.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 528 RSGKPVRV 535  
|||||  
Db 193 RSGKPVRV 200

RESULT 6  
H97324  
uncharacterized protein, homolog of B. anthracis (gi\_48942631) [imported] - Clostridium  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: H97324  
R;Nolling, J.; Berton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: H97324  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-265 <KUR>  
A;Cross-references: UNIPROT:Q97DL6; GB:AE001437; PIDN:AAK81387.1; PTD:gl5026549; GSPDB:G  
A;Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:  
A:Gene: CAC3458

Query Match 1.0%; Score 8; DB 2; Length 265;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 EKKGSGFL 574  
|||||  
Db 248 EKKGSGFL 255

## RESULT 7

S23561  
HAL1 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein LP25c; protein YP9723.05c; protein YPR005c  
C:Species: Saccharomyces cerevisiae  
C>Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 09-Jul-2004  
C:Accession: S23561; S52818; S59750  
R:Gaxiola, R.; de Larinoa, I.F.; Villalba, J.M.; Serrano, R.  
EMBO J. 11, 3157-3164, 1992  
A>Title: A novel and conserved salt-induced protein is an important determinant of salt  
A:Reference number: S23561; MUID:92371421; PMID:1505513  
A:Accession: S23561  
A:Molecule type: DNA  
A:Residues: 1-294 <GAX>  
A:Cross-references: UNIPROT:Q01766; EMBL:X67559; NID:g3760; PIDN:CAA47858.1; PID:g3761  
A:Experimental source: strain RS-16  
R:Pearson, D.; Bowman, S.  
submitted to the EMBL Data Library, April 1995  
A:Reference number: S52814  
A:Accession: S52818  
A:Molecule type: DNA

A:Residues: 1-294 <PEA>  
A:Cross-references: EMBL:Z48951; NID:g762999; PIDN:CAA88783.1; PID:g763004; MIPS:YPR005c  
A:Experimental source: strain AB972  
R:Wang, Y.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.K.; VC  
submitted to the EMBL Data Library, July 1995  
A>Description: The sequence of Saccharomyces cerevisiae chromosome XVI right arm.  
A:Reference number: S59746  
A:Accession: S59750  
A:Molecule type: DNA  
A:Residues: 1-294 <WAN>  
A:Cross-references: EMBL:U31900; NID:g1276597; PIDN:AAA97584.1; PID:g939739; MIPS:YPR005c  
C:Genetics:  
A:Gene: SGD:HAL1  
A:Cross-references: SGD:S0006209; MIPS:YPR005c  
A:Map position: 16R  
C:Superfamily: Saccharomyces cerevisiae HAL1 protein

Query Match 1.0%; Score 8; DB 2; Length 294;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 476 FTYTSGGG 483  
|||||  
Db 81 FTYTSGGG 88

## RESULT 8

A82406  
Cytochrome d ubiquinol oxidase, chain II VCA0873 [imported] - Vibrio cholerae (strain NI  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: A82406  
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
l. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: A82406  
A>Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-378 <HEI>

A:Cross-references: UNIPROT:Q9KL74; GB:AE004415; GB:AE003853; NID:g9658303; PIDN:AAF96773  
A:Experimental source: serogroup O1; strain NI6961; biotype El Tor  
C:Genetics:  
A:Gene: VCA0873  
A:Map position: 2  
C:Superfamily: cytochrome d ubiquinol oxidase

Query Match 1.0%; Score 8; DB 2; Length 378;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 LQQSSESDK 106  
|||||  
Db 39 LQQSSESDK 46

## RESULT 9

T26968  
Hypothetical protein Y47H9C.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T26968  
R:Harris, B.  
submitted to the EMBL Data Library, October 1998

A:Reference number: Z20293  
A:Accession: T26968  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-394 <WIL>  
A:Cross-references: UNIPROT:Q9XWE0; EMBL:AL032657; PIDN:CAA21735.1; GSPDB:GN00019; CESP:1  
A:Experimental source: clone Y47H9C  
C:Genetics:  
A:Gene: CESP:Y47H9C.6  
A:Map position: 1  
A:Introns: 31/3; 46/1; 76/1; 92/1; 118/1; 156/1; 194/2; 287/3; 332/3; 347/1; 395/1

Query Match 1.0%; Score 8; DB 2; Length 394;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 649 RKSAGGGP 656  
|||||  
Db 41 RKSAGGGP 48

## RESULT 10

S66713  
Hypothetical protein YOL030w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein O2145  
C:Species: Saccharomyces cerevisiae  
C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: S66713  
R:Habbig, B.; Hattenhorst, U.; Hollenberg, C.P.; Ramezani Rad, M.  
submitted to the Protein Sequence database, July 1996  
A:Reference number: S66703  
A:Accession: S66713  
A:Molecule type: DNA  
A:Residues: 1-484 <HAB>

A:Cross-references: UNIPROT:Q08193; EMBL:Z74772; NID:g1419818; PIDN:CAA99030.1; PID:g1419  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: MIPS:YOL030w  
A:Cross-references: SGD:S0005390  
A:Map position: 15L  
C:Superfamily: glycopospholipid-anchored surface glycoprotein GAS1

Query Match 1.0%; Score 8; DB 2; Length 484;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 ASATSSSQ 394



```

Db      448 ASATSSSQ 455
|||||
RESULT 11
S66763
hypothetical protein YOL070c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein Ol150
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: S66763
R;Alexandraki, D.; Katsoulou, C.; Tzermia, M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66756
A;Accession: S66763
A;Molecule type: DNA
A;Residues: 1-501 <ALE>
A;Cross-references: UNIPROT:Q08229; EMBL:Z74812; NID:g1419894; PID:e251874; PID:g1419895
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YOL070C
A;Cross-references: SGD:S0005431
A;Map position: 15L

Query Match      1.0%; Score 8; DB 2; Length 501;
Best Local Similarity 100.0%; Pred. No. 18;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      16 VDSLSRLT 23
Db      467 VDSLSRLT 474
|||||

RESULT 12
T21134
hypothetical protein F20C5.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T21134
R;Matthews, P.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19381
A;Accession: T21134
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1130 <WIL>
A;Cross-references: UNIPROT:Q8MQ68; EMBL:Z68161; PIDN:CAA92295.1; GSPDB:GN00022; CESP:F2
A;Experimental source: clone F20C5
C;Genetics:
A;Gene: CESP:F20C5.2
A;Map position: 4
A;Introns: 15/3; 34/2; 69/2; 152/3; 183/3; 207/2; 314/3; 513/3; 546/2; 594/1; 632/3; 667/3

Query Match      1.0%; Score 8; DB 2; Length 1130;
Best Local Similarity 100.0%; Pred. No. 36;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      387 ASATSSSQ 394
Db      959 ASATSSSQ 966
|||||

RESULT 13
G86467
hypothetical protein F7P12.4 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G86467
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

```

```

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86467
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1587 <STO>
A;Cross-references: UNIPROT:Q9C8N3; GB:AE005172; NID:g10092383; PIDN:AAG12790.1; GSPDB:G
C;Genetics:
A;Map position: 1

Query Match      1.0%; Score 8; DB 2; Length 1587;
Best Local Similarity 100.0%; Pred. No. 48;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      350 SSVPSSEDE 357
Db      319 SSVPSSEDE 326
|||||

RESULT 14
S03778
uviB protein - Clostridium perfringens plasmid PIP404
C;Species: Clostridium perfringens
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S03778; J0354
R;Garnier, T.; Cole, S.T.
Mol. Microbiol. 2, 607-614, 1988
A;Title: Studies of UV-inducible promoters from Clostridium perfringens in vivo and in vi
A;Reference number: S03777; MUID:89039249; PMID:2460717
A;Accession: S03778
A;Molecule type: DNA
A;Residues: 1-64 <GAR>
A;Cross-references: UNIPROT:P15936; EMBL:M32882; NID:g150738; PIDN:AAA98258.1; PID:g1507
C;Genetics:
A;Gene: uviB
A;Genome: plasmid
C;Superfamily: uviB protein

Query Match      0.9%; Score 7; DB 1; Length 64;
Best Local Similarity 100.0%; Pred. No. 31;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      622 KENSKRE 628
Db      28 KENSKRE 34
|||||

RESULT 15
T42941
hypothetical protein - ateline herpesvirus 3 (strain 73)
C;Species: ateline herpesvirus 3
A;Variety: strain 73
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42941
R;Albrecht, J.C.; Fleckenstein, B.
submitted to the EMBL Data Library, August 1998
A;Description: Primary structure of the herpesvirus ateles genome.
A;Reference number: Z22274
A;Accession: T42941
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-74 <ALB>
A;Cross-references: UNIPROT:Q9YTN8; EMBL:AF083424; PIDN:AAC95552.1
A;Experimental source: strain 73

Query Match      0.9%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 36;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



QY 557 GYKVK 563  
| | | | |  
Db 38 GYKVK 44

Search completed: November 1, 2004, 15:55:41  
Job time : 30 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2004, 15:45:53 ; Search time 92 Seconds  
(without alignments)

3092.091 Million cell updates/sec

Title: US-10-019-071-2

Perfect score: 793

Sequence: 1 MWIQRVTMDGRQTHVTDSLS.....VNQPLQTVLNQLFPGYGNR 793

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	793	100.0	793	4	Abb76980 Human Inv
2	591	74.5	793	6	Abz48157 Human bla
3	591	74.5	793	7	Adf61820 Human NP9
4	591	74.5	793	7	Adf76781 Novel hum
5	591	74.5	793	8	Adn05229 Antipsori
6	591	74.5	793	8	Ado20357 Human PRO
7	578	72.9	780	6	Abu56628 Lung canc
8	578	72.9	780	7	Adn38838 Cancer/an
9	174	21.9	174	4	Abb76983 Human Inv
10	86	10.8	133	4	Aau16348 Human nov
11	86	10.8	133	6	Abu55417 Human nov
12	80	10.1	148	4	Adm20128 Protein e
13	80	10.1	150	4	Adm19888 Protein e
14	25	3.2	26	4	Abb76982 Human Inv
15	19	2.4	198	5	Abp64013 Human ORF
16	19	2.4	645	6	Abu69599 Human NF-
17	19	2.4	645	6	Abu69620 Human NF-
18	19	2.4	802	6	Abz82238 Human nuc
19	15	1.9	15	4	Abb76981 Human Inv
20	14	1.8	617	7	Adb30490 Plant yie
21	14	1.8	617	8	Adi43851 Plant tra
22	13	1.6	87	5	Abb97911 Human sec
23	10	1.3	71	4	Aam40216 Human pol
24	10	1.3	71	4	Aab73737 Ring fing
25	10	1.3	110	4	Aau15895 Human nov

26	10	1.3	110	6	ABU54964	Abu54964 Human nov
27	10	1.3	178	4	AAU16350	Aau16350 Human nov
28	10	1.3	178	6	ABU55419	Abu55419 Human nov
29	10	1.3	180	4	AAU42002	Aam42002 Human pol
30	9	1.1	92	4	AAU76581	Aag76581 Human col
31	9	1.1	422	7	ABO71837	AbO71837 Pseudomon
32	9	1.1	1195	6	ABU11484	Abu11484 Human MDD
33	9	1.1	1233	5	ABP55147	Abp55147 Neurodeve
34	9	1.1	1240	6	ABU54572	Abu54572 Human NOV
35	9	1.1	1279	7	ADI21252	Adi21252 Novel hum
36	9	1.1	1284	7	ADJ69202	Adj69202 Human hea
37	8	1.0	48	4	ABB6651	Abb6651 Drosophil
38	8	1.0	52	4	AAU49432	Aau49432 Propionib
39	8	1.0	52	6	ABM45951	Abm45951 Propionib
40	8	1.0	56	4	AAU40628	Aau40628 Propionib
41	8	1.0	56	6	ABM37147	Abm37147 Propionib
42	8	1.0	105	4	ABB66912	Abb66912 Drosophil
43	8	1.0	129	4	AAU48329	Aau48329 Propionib
44	8	1.0	129	6	ABM44848	Abm44848 Propionib
45	8	1.0	166	5	ABP32211	Abp32211 Human ORF

## ALIGNMENTS

### RESULT 1

ABB76980  
ID ABB76980 standard; protein; 793 AA.

XX AC ABB76980;

XX DT 22-JUL-2002 (first entry)

XX DE Human Inverted CCAAT box binding protein, ICBP90.

XX KW Human; inverted CCAAT box binding protein; ICBP90; cytostatic;  
cell proliferation control; inverted CCAAT box; cancer.

XX OS Homo sapiens.

XX PN WO200078949-A1.

XX PD 28-DEC-2000.

XX PF 22-JUN-2000; 2000WO-FR001747.

XX PR 22-JUN-1999; 99FR-00007935.

XX (ADER-) ADEREGEM ASSOC DEV RECH EN GENETIQ.

XX Bronner C, Hopfner R, Mousli M, Jeltsch J, Lutz Y, Oudet P;

WPI; 2001-091571/10.

XX N-PSDB; ABL58020.

XX Novel inverted CCAAT box binding protein, and related nucleic acids,  
antibodies and specific ligands, useful for treating and preventing  
cancer.

XX Claim 1; Fig 7; 115pp; French.

XX The present sequence is the protein sequence for human ICBP90 (inverted  
CCAAT box binding protein). The inverted CCAAT box is implicated in cell  
proliferation control. Several copies of the inverted CCAAT box are  
present in the promoter of the topoisomerase IIalpha gene, and also  
functions as a nuclear receptor. ICBP90 and its coding sequence are  
useful for treatment and/or prevention of cancer

XX SQ Sequence 793 AA;

Query Match

Best Local Similarity 100.0%; Score 793; DB 4; Length 793;

Matches 793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy	421	IPGIPVGTMMFRFVQVSSGSHRPHVAGIHGRSNDGYSYLVLAGGYEDDVDRHGNFFTYTG	480
Db	421	IPGIPVGTMMFRFVQVSSGSHRPHVAGIHGRSNDGYSYLVLAGGYEDDVDRHGNFFTYTG	480
Qy	481	SGGRDLSGNKRTABQSCDOKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRVNVK	540
Db	481	SGGRDLSGNKRTABQSCDOKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVVRVNVK	540
Qy	541	GGKNSKYAPAEGRVDGIYKVVKYWPXGSGFLVWRYLLRRDDDEPGPWTKEGDKRIKK	600
Db	541	GGKNSKYAPAEGRVDGIYKVVKYWPXGSGFLVWRYLLRRDDDEPGPWTKEGDKRIKK	600
Qy	601	LGLTWOYPEGYLEALANRERKENSKEEEEEQOEGGFASPRTGKWKWRKKSAGGSPSRAG	660
Db	601	LGLTWOYPEGYLEALANRERKENSKEEEEEQOEGGFASPRTGKWKWRKKSAGGSPSRAG	660
Qy	661	SPRSTSKTKVEPSYLTAAQSSLIREDKSNAKLWNEVLASLKDRASSGPPOLFYSKVEE	720
Db	661	SPRSTSKTKVEPSYLTAAQSSLIREDKSNAKLWNEVLASLKDRASSGPPOLFYSKVEE	720
Qy	721	TFQICICQOELVFRPIITTVCOHNVCKDLDRAFQVFCSPACRYDLGRSYAMQVNOPLQT	780
Db	721	TFQICICQOELVFRPIITTVCOHNVCKDLDRAFQVFCSPACRYDLGRSYAMQVNOPLQT	780
Qy	781	VLNOLFFPGYGNGR	793
Db	781	VLNOLFFPGYGNGR	793

## RESULT. T 3

RESOL 3  
ADF61820

ID ADF61820 standard: protein: 793 AA.

AD 01020  
XX  
AC ADF61820:

XX  
DT 12-FEB-2004 (first entry)

XX  
DE Human NP95 pr

CC within a cell with a compound and determining the chemical or phenotypic  
CC effect of the compound upon the cell. The method of the invention has  
CC cytostatic, antiproliferative, antiatherosclerotic, vasotropic and  
CC antithrombotic applications and may be useful for identifying a compound  
CC that modulates cell cycle arrest. Such compounds may subsequently be used  
CC for developing therapeutic reagents to treat melanoma, breast, ovarian,  
CC lung, gastrointestinal or colon cancer, as well as other proliferative  
CC diseases such as Grave's disease, psoriasis, atherosclerosis, restenosis  
CC and other vasoproliferative diseases. The current sequence is that of the  
CC human NP95 protein of the invention.

RESULT 4	QY	61	YEVRINDTTQLLVROSLVPHSTKERDSELSDDTSGCCLGQSDSKSTHGEAAETDSR	120
ADF76781	Db	61	YEVRINDTTQLLVROSLVPHSTKERDSELSDDTSGCCLGQSDSKSTHGEAAETDSR	120
XX	QY	121	PADEDMWDETELGLYKVNVEYVDARDTNMGAWFAQVVRVTRKAPSRDEPCSSTSRPALEE	180
AC	Db	121	PADEDMWDETELGLYKVNVEYVDARDTNMGAWFAQVVRVTRKAPSRDEPCSSTSRPALEE	180
XX	QY	181	DVIYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNLNPNPKERGFWYD	240
DT	Db	181	DVIYHVKYDDYPENGVMNSRDVRARARTIIKWQDLEVGQVVMNLNPNPKERGFWYD	240
XX	QY	241	AEISRKRETTARELVANVVLGDDSLNDCRIIFVDEVFKIERPEGSEPMVNDPMRRKSGP	300
XX	Db	241	AEISRKRETTARELVANVVLGDDSLNDCRIIFVDEVFKIERPEGSEPMVNDPMRRKSGP	300
XX	QY	301	SKHKCDDVNRLCRVCACHLGGGRQDPDKOLMCDCECDMAFHYYCLDPLSPSEDEWYC	360
XX	Db	301	SKHKCDDVNRLCRVCACHLGGGRQDPDKOLMCDCECDMAFHYYCLDPLSPSEDEWYC	360
XX	QY	361	PECRNDASEVVLAGERLRESKKNAKMASATSSORDWKGKMACVGRTEKCTIIVPSNHYGP	420
XX	Db	361	PECRNDASEVVLAGERLRESKKNAKMASATSSORDWKGKMACVGRTEKCTIIVPSNHYGP	420
XX	QY	421	IPGIPVGTMRFRVQVSESGVHRPHVAGIHGRSNDGSYSVLVAGGYEDDDVHGNNFFTYTG	480
XX	Db	421	IPGIPVGTMRFRVQVSESGVHRPHVAGIHGRSNDGSYSVLVAGGYEDDDVHGNNFFTYTG	480
XX	QY	481	SGGRDLSGNKRRTABQSCDQKLTNTNRALALNCAPINDQEGAEAKDWRSGKPVVRVNVK	540
XX	Db	481	SGGRDLSGNKRRTABQSCDQKLTNTNRALALNCAPINDQEGAEAKDWRSGKPVVRVNVK	540
XX	QY	541	GGKNSKYAPAEGRNYDGIYKVKYWPBKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKK	600
XX	Db	541	GGKNSKYAPAEGRNYDGIYKVKYWPBKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKK	600
XX	QY	601	LGLTMQYPEGVLEALANREKENSKEEBEQGGFASPRGTGKWKRSAGGSPSAG	660
XX	Db	601	LGLTMQYPEGVLEALANREKENSKEEBEQGGFASPRGTGKWKRSAGGSPSAG	660
XX	QY	661	SPRSTSKTKVPEYSLTAQSSLIREDKSNAKLWNEVLASLKDRPASGSPQLFLSKVEE	720
XX	Db	661	SPRSTSKTKVPEYSLTAQSSLIREDKSNAKLWNEVLASLKDRPASGSPQLFLSKVEE	720
XX	QY	721	TFQCICCOELVFRITTVCOHNVCKDCLDSFRAQVSCPCRYDLGRSYAMVNOPLQT	780
XX	Db	721	TFQCICCOELVFRITTVCOHNVCKDCLDSFRAQVSCPCRYDLGRSYAMVNOPLQT	780
XX	QY	781	VLNOLFFGYNGR 793	
XX	Db	781	VLNOLFFGYNGR 793	
XX	QY	793	ADN05229 standard; protein; 793 AA.	
XX	Db	793	ADN05229 standard; protein; 793 AA.	
XX	QY	793	ADN05229;	
XX	Db	793	ADN05229;	
XX	QY	793	01-JUL-2004 (first entry)	
XX	Db	793	01-JUL-2004 (first entry)	
XX	QY	793	Antipsoriatic protein sequence #791.	
XX	Db	793	Antipsoriatic protein sequence #791.	
XX	QY	793	antipsoriatic; gene therapy; psoriasis; diagnosis.	
XX	Db	793	antipsoriatic; gene therapy; psoriasis; diagnosis.	
XX	QY	793	Homo sapiens.	
XX	Db	793	Homo sapiens.	
XX	QY	793	WO2004028479-A2.	
XX	Db	793	WO2004028479-A2.	
XX	QY	793	08-APR-2004.	
XX	Db	793	08-APR-2004.	
XX	QY	793	25-SEP-2003; 2003WO-US030907.	
XX	Db	793	25-SEP-2003; 2003WO-US030907.	

Novel human secreted and transmembrane protein SeqID 456.

human; PRO; membrane bound protein; membrane bound receptor; cell proliferation; cell migration; cell differentiation; mitogenic factor; survival factor; cytotoxic factor; differentiation factor; neuroepitope; hormone; cell receptor; receptor-ligand interaction; cytostatic; chondrocyte; tumour.

Homo sapiens.

WO2003072035-A2.

04-SEP-2003.

21-FEB-2003; 2003WO-US005241.

22-FEB-2002; 2002US-0359461P.

(GETH ) GENENTECH INC.

Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR; Williams PM, Wood WI, Wu TD; WPI; 2003-721702/68.

N-PSDB; ADF76780.

New PRO polypeptides, useful for diagnosing and treating an immune related disorder, e.g. systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or diabetes mellitus.

Claim 10; SEQ ID NO 456; 918pp; English.

This invention relates to novel nucleic acids encoding human PRO secreted and transmembrane proteins. Extracellular proteins play important roles in the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells (for example proliferation, migration or differentiation) is typically governed by information received from other cells and the immediate environment. The information is often transmitted by secreted polypeptides (for example mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuroepitopes and hormones) which are received and interpreted by diverse cell receptors or membrane bound proteins. These membrane bound proteins and receptors may be of use as pharmaceutical and diagnostic agents, such as in the blocking of receptor-ligand interactions. The current invention provides the amino acid sequences of novel human membrane bound receptors and proteins, along with the cDNA sequences encoding them. The novel proteins of the invention may have cytostatic activities through the stimulation of chondrocytes. The nucleic acids of the invention may be useful for the manufacture of a medicament for diagnosing or treating a tumour in a mammal. In addition, they may be useful for measuring or detecting the expression of a tumour associated gene. The present sequence is the amino acid sequence of a human PRO protein of the invention.

Query Match 74.5%; Score 591; DB 7; Length 793; Best Local Similarity 99.7%; Pred. No. 0; Matches 791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MWIQVTRMDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60

Db 1 MWIQVTRMDGRQTHTVDSLSRLTKVEELRRKIQELFHVPEGLQRLFYRGKQMEDGHTLFD 60



```
Best Local Similarity 99.7%; Pred. No. 0;
Matches 791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MWIQRVMDGRQTHVDSLSRLTKVEELRRKIQLFHVPEGLQRLFYRGKQMEDGHTLFD 60
Db 1 MWIQRVMDGRQTHVDSLSRLTKVEELRRKIQLFHVPEGLQRLFYRGKQMEDGHTLFD 60

QY 61 YEVRINDTIQLLVRSQSLVPHSTKERDSELSDTSGCCLGQSESDKSTHGEAAETDSR 120
Db 61 YEVRINDTIQLLVRSQSLVPHSTKERDSELSDTSGCCLGQSESDKSTHGEAAETDSR 120

QY 121 PAEDMDWDETELGLYKVNVEYDARDTNNMGAFEAQVVRVTRKAPSRDPCSTSRPALEE 180
Db 121 PAEDMDWDETELGLYKVNVEYDARDTNNMGAFEAQVVRVTRKAPSRDPCSTSRPALEE 180

QY 181 DVIHVKYDDYPENGVOVMSNRDVRARARTIIKWQDLEVGQVVMNPNPKRGFWYD 240
Db 181 DVIHVKYDDYPENGVOVMSNRDVRARARTIIKWQDLEVGQVVMNPNPKRGFWYD 240

QY 241 ABEISRKRETRARELYANVLGDDSLNDCRIIFVDEVEFKIERPBGSGPMVNDPMRKGSP 300
Db 241 ABEISRKRETRARELYANVLGDDSLNDCRIIFVDEVEFKIERPBGSGPMVNDPMRKGSP 300

QY 301 SCCKCKODVNLRCRYCACHLCGRQDPDKQLMCDCEDMAFHIYCLDPLSSVPSEDEWYC 360
Db 301 SCCKCKODVNLRCRYCACHLCGRQDPDKQLMCDCEDMAFHIYCLDPLSSVPSEDEWYC 360

QY 361 PECDNDASEVLAGERLESKNAMASATSSQBDWKGKMACVGRTECTIVPSNHYGP 420
Db 361 PECDNDASEVLAGERLESKNAMASATSSQBDWKGKMACVGRTECTIVPSNHYGP 420

QY 421 IPGIPVGTWFRVQVSGVHRPHVAGIHGKSNDSYSLVLAGGYDDVDHGNPFTYTG 480
Db 421 IPGIPVGTWFRVQVSGVHRPHVAGIHGKSNDSYSLVLAGGYDDVDHGNPFTYTG 480

QY 481 SGGDLGSKNRTAEQSCDQKLTNTNRALALNCFAPIINDQEGAEAKDWSGKPRVVRNVK 540
Db 481 SGGDLGSKNRTAEQSCDQKLTNTNRALALNCFAPIINDQEGAEAKDWSGKPRVVRNVK 540

QY 541 GKKSKYAPAGNRRVDGIIKVKVWPEKSGFLVWRVLLRRDDDEPGWTKGKDR1KK 600
Db 541 GKKSKYAPAGNRRVDGIIKVKVWPEKSGFLVWRVLLRRDDDEPGWTKGKDR1KK 600

QY 601 LGLTMQYEGYLEALANERERKENS KREBEQEGGFASPRTGKWKRSAGGSPSRAG 660
Db 601 LGLTMQYEGYLEALANERERKENS KREBEQEGGFASPRTGKWKRSAGGSPSRAG 660

QY 661 SPRTSKTKVPEYSLTAQOSLIREDKSNAKLMNEVLASLKDRPASGSPQLFLSKVEE 720
Db 661 SPRTSKTKVPEYSLTAQOSLIREDKSNAKLMNEVLASLKDRPASGSPQLFLSKVEE 720

QY 721 TFQCIICCOELVFRPITTVCOHNVCCKDLDRSPRAQVFCPCRYDLGRSYAMQVNPLOT 780
Db 721 TFQCIICCOELVFRPITTVCOHNVCCKDLDRSPRAQVFCPCRYDLGRSYAMQVNPLOT 780

QY 781 VLNLQFPYGNNGR 793
Db 781 VLNLQFPYGNNGR 793

RESULT 7
ID ABUS56628 standard; protein; 780 AA.
XX
AC ABUS56628;
XX
DT 02-APR-2003 (first entry)
DE Lung cancer-associated polypeptide #221.
XX
KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
```

```
chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

Unidentified.
WO200286443-A2.
31-OCT-2002.
18-APR-2002; 2002WO-US012476.
18-APR-2001; 2001US-0284770P.
10-MAY-2001; 2001US-0290492P.
09-NOV-2001; 2001US-0339245P.
13-NOV-2001; 2001US-0350666P.
29-NOV-2001; 2001US-0334370P.
12-APR-2002; 2002US-0372246P.
(EOSB-) EOS BIOTECHNOLOGY INC.
Aziz N, Murray R;
WPI; 2003-093161/08.
N-PSDB; ABX76357.
Detecting a lung cancer-associated transcript in a cell from a patient
for treating lung cancer, by contacting a biological sample from the
patient with a polynucleotide that exhibits increased or decreased
expression in lung cancer.
Claim 27; Page 357; 453pp; English.
The invention relates to a method for detecting a lung cancer-associated
transcript in a cell from a patient, comprising contacting a biological
sample from the patient with a polynucleotide that selectively hybridises
to a sequence that is at least 80 % identical to a gene that exhibits
increased or decreased expression in lung cancer samples. Lung cancer-
associated polynucleotides and polypeptides are used for identifying a
compound that modulates a lung cancer-associated polypeptide, for
inhibiting proliferation of a lung cancer-associated cell to treat lung
cancer in a patient and for treating a mammal having lung cancer by
administering a modulatory compound identified. The methods are useful
for treating lung cancer, such as small cell lung cancer, non-small cell
lung cancer or other benign or precancerous lesions, e.g. atelectasis,
emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
bronchiectasis. The genes, polynucleotides and polypeptides are useful
for diagnostic purposes and as targets for screening for therapeutic
compounds that modulate lung cancer, such as antibodies. Sequences
ABUS56408-ABUS56745 represent lung cancer-associated polypeptides of the
invention.
Sequence 780 AA;
Query Match 72.9%; Score 578; DB 6; Length 780;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 778; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MWIQRVMDGRQTHVDSLSRLTKVEELRRKIQLFHVPEGLQRLFYRGKQMEDGHTLFD 60
Db 1 MWIQRVMDGRQTHVDSLSRLTKVEELRRKIQLFHVPEGLQRLFYRGKQMEDGHTLFD 60

QY 61 YEVRINDTIQLLVRSQSLVPHSTKERDSELSDTSGCCLGQSESDKSTHGEAAETDSR 120
Db 61 YEVRINDTIQLLVRSQSLVPHSTKERDSELSDTSGCCLGQSESDKSTHGEAAETDSR 120

QY 121 PAEDMDWDETELGLYKVNVEYDARDTNNMGAFEAQVVRVTRKAPSRDPCSTSRPALEE 180
Db 121 PAEDMDWDETELGLYKVNVEYDARDTNNMGAFEAQVVRVTRKAPSRDPCSTSRPALEE 180

QY 181 DVIHVKYDDYPENGVOVMSNRDVRARARTIIKWQDLEVGQVVMNPNPKRGFWYD 240
Db 181 DVIHVKYDDYPENGVOVMSNRDVRARARTIIKWQDLEVGQVVMNPNPKRGFWYD 240
```



```
QY 241 AHSKRTRTRARELYANVVLGDDSLNDCRIIFVDEVFKIERPBGSPMVDNPMRKSGP 300
D 241 AHSKRTRTRARELYANVVLGDDSLNDCRIIFVDEVFKIERPBGSPMVDNPMRKSGP 300
QY 301 SCKHCKDDVNRLCRVACACCHLGGRRDPDKQLMCDCECDMAFHLYCLDPLSSVPSSEDEWYC 360
D 301 SCKHCKDDVNRLCRVACACCHLGGRRDPDKQLMCDCECDMAFHLYCLDPLSSVPSSEDEWYC 360
QY 361 PECRNDASEVVLAGERLRESKKNKAKMASATSSSQDRDWGKMACVGRTECTIVPSNHGYP 420
D 361 PECRNDASEVVLAGERLRESKKNKAKMASATSSSQDRDWGKMACVGRTECTIVPSNHGYP 420
QY 421 IFGIPVGTWRRFRVQVSEGVHRPHVAGIHGRSNDGSLYLAGYEDVDHGNFTYTG 480
D 421 IFGIPVGTWRRFRVQVSEGVHRPHVAGIHGRSNDGSLYLAGYEDVDHGNFTYTG 480
QY 481 SGRDLSGNKRTAEGSCDQKLTNTRALNCFAPINDOEGAEAKDWRSGKPRVVRNVK 540
D 481 SGRDLSGNKRTAEGSCDQKLTNTRALNCFAPINDOEGAEAKDWRSGKPRVVRNVK 540
QY 541 GGNKSKYAPAEGRNRYDGIYKVYKWPCKGSGFLVRYLLRRDDDEPGWTKEGDKRIKK 600
D 541 GGNKSKYAPAEGRNRYDGIYKVYKWPCKGSGFLVRYLLRRDDDEPGWTKEGDKRIKK 600
QY 601 LGLTWQYPEGYLEALANRREKENSREBEEOEGGFASPTGKGKWKKSAGGSPSRAG 660
D 601 LGLTWQYPEGYLEALANRREKENSREBEEOEGGFASPTGKGKWKKSAGGSPSRAG 660
QY 661 SPRTSKTKVBPYSYLTQOSSLIREDKSNALWNEVLASLAKDRPASGSPFOLFJSKVEE 720
D 661 SPRTSKTKVBPYSYLTQOSSLIREDKSNALWNEVLASLAKDRPASGSPFOLFJSKVEE 720
QY 721 TPQCCICQELVFRPITTVQHNVCCKDLRSFRAQVFCPACRYDILGRSYAMQVNPLOT 780
D 721 TPQCCICQELVFRPITTVQHNVCCKDLRSFRAQVFCPACRYDILGRSYAMQVNPLOT 780
RESULT 8
ADN38838
ID ADN38838 standard; protein; 780 AA.
XX
AC ADN38838;
XX
DT 17-JUN-2004 (first entry)
DE
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:156.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-032464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
```

```
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;
XX
XX WPI; 2003-468649/44.
XX N-PSDB; ADN38837.
XX
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX
XX Claim 12; SEQ ID NO 156; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or
XX other diseases such as angiogenic or fibrotic disorders, and to methods
XX of determining the presence or absence of a pathological cell in a
XX patient by detecting a nucleic acid at least 80% identical to those of
XX the invention or by detecting a polypeptide of the invention. The
XX invention also relates to expression vectors and host cells comprising a
XX nucleic acid of the invention; antibodies which specifically bind a
XX polypeptide of the invention; use of such antibodies for drug targeting;
XX and methods of screening for modulators of activity or expression of the
XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX antibodies and methods are useful for diagnosing, prognosing and treating
XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX neovascularisation syndromes, scarring and uterine fibroids. They may
XX also be useful in wound healing and in contraception. The present
XX sequence represents a polypeptide of the invention.
XX
XX Sequence 780 AA;
XX
XX Query Match 72.9%; Score 578; DB 7; Length 780;
XX Best Local Similarity 99.7%; Pred. No. 0;
XX Matches 778; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MWIQRVTWDRQTHVDSLSRLTKVEELRRKTOELFHVPEGLQRLFYRGKQMEDGHTLFD 60
D 1 MWIQRVTWDRQTHVDSLSRLTKVEELRRKTOELFHVPEGLQRLFYRGKQMEDGHTLFD 60
QY 61 YEVRNDTIQLVRSQSLVPLPHSTKERDSELSDTSGCCLGQSESKSSTHGAATAETDSR 120
D 61 YEVRNDTIQLVRSQSLVPLPHSTKERDSELSDTSGCCLGQSESKSSTHGAATAETDSR 120
QY 121 PADEDMDETELGLYKVNVEYVDARTNMGANPEAQVVRTRKAPSRDEPCSTSRPALEE 180
D 121 PADEDMDETELGLYKVNVEYVDARTNMGANPEAQVVRTRKAPSRDEPCSTSRPALEE 180
QY 181 DVIYHVKYDDYPENGVOVMSNDVBARARTIWKQDLEVGQVVMNPNPNKRGFWYD 240
D 181 DVIYHVKYDDYPENGVOVMSNDVBARARTIWKQDLEVGQVVMNPNPNKRGFWYD 240
QY 241 AEISKRTRTRARELYANVVLGDDSLNDCRIIFVDEVFKIERPBGSPMVDNPMRKSGP 300
D 241 AEISKRTRTRARELYANVVLGDDSLNDCRIIFVDEVFKIERPBGSPMVDNPMRKSGP 300
QY 301 SCKHCKDDVNRLCRVACACCHLGGRRDPDKQLMCDCECDMAFHLYCLDPLSSVPSSEDEWYC 360
D 301 SCKHCKDDVNRLCRVACACCHLGGRRDPDKQLMCDCECDMAFHLYCLDPLSSVPSSEDEWYC 360
QY 361 PECRNDASEVVLAGERLRESKKNKAKMASATSSSQDRDWGKMACVGRTECTIVPSNHGYP 420
```

Db 361 PECDNDASEVLAGERLRESKKAKMASATSSSQEDWKGWACVGRTECTIVSNHGP 420  
QY 421 IPGIPVGTWFRVQVSSGVHRPHVAGIHGRSNDGSYSLVLAGYEDDVHGNFFVTG 480  
Db 421 IPGIPVGTWFRVQVSSGVHRPHVAGIHGRSNDGAYSLVLAGYEDDVHGNFFVTG 480  
QY 481 SGGDLSCNKTAEQSCDKLTNTNRALALNCFAPINDOEGAEKDWESGKPVVRNVK 540  
Db 481 SGGDLSCNKTAEQSCDKLTNTNRALALNCFAPINDOEGAEKDWESGKPVVRNVK 540  
QY 541 GGKNSKYAPAEGRNRYDGIYKVKWPEKSGFLVWRYLLRRDDDEPGPWTKEGDKRIKK 600  
Db 541 GGKNSKYAPAEGRNRYDGIYKVKWPEKSGFLVWRYLLRRDDDEPGPWTKEGDKRIKK 600  
QY 601 LGLTMQYPEGYLEALANREREKENSREBEEOQEGGFASPRGTGKWKRSAGGSPSAG 660  
Db 601 LGLTMQYPEGYLEALANREREKENSREBEEOQEGGFASPRGTGKWKRSAGGSPSAG 660  
QY 661 SPRTSKTKTVEPYSLTAQSSLIREDKSNKLNWNEVLASLKDPRPASGSPQLFLSKVEE 720  
Db 661 SPRTSKTKTVEPYSLTAQSSLIREDKSNKLNWNEVLASLKDPRPASGSPQLFLSKVEE 720  
QY 721 TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCRYDLGRSYAMQVNPQLQT 780  
Db 721 TFQCICCOELVFRPITTVCOHNVCKDCLDRSFRAQVFCPCRYDLGRSYAMQVNPQLQT 780  
RESULT 9  
ABB76983  
ID ABB76983 standard; protein; 174 AA.  
XX  
AC ABB76983;  
XX  
DT 22-JUL-2002 (first entry)  
XX  
DE Human Inverted CCAAT box binding protein, ICBP90, fragment #3.  
XX  
KW Human; inverted CCAAT box binding protein; ICBP90; cytostatic;  
KW cell proliferation control; inverted CCAAT box; cancer.  
XX  
OS Homo sapiens.  
XX  
FN WO200078949-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000MO-FR001747.  
XX  
PR 22-JUN-1999; 99FR-00007935.  
XX  
PA (ADER-) ADEREGEM ASSOC DEV RECH EN GENETIQ.  
XX  
PI Bronner C, Hopfner R, Mousli M, Jeltsch J, Lutz Y, Oudet P;  
XX WPI; 2001-091571/10.  
XX N-PSDB; ABL58023.  
XX  
PT Novel inverted CCAAT box binding protein, and related nucleic acids,  
PT antibodies and specific ligands, useful for treating and preventing  
PT cancer.  
XX  
PS Claim 2; Page 103; 115pp; French.  
XX  
CC The present sequence is a protein fragment of human ICBP90 (inverted  
CC CCAAT box binding protein). The inverted CCAAT box is implicated in cell  
CC proliferation control. Several copies of the inverted CCAAT box are  
CC present in the promoter of the topoisomerase I alpha gene, and also  
CC functions as a nuclear receptor. ICBP90 (ABB76980) and its coding  
CC sequence (ABL58020) are useful for treatment and/or prevention of cancer  
XX  
SQ Sequence 174 AA;

Query Match 21.9%; Score 174; DB 4; Length 174;  
Best Local Similarity 100.0%; Pred. No. 4.4e-170; Indels 0; Gaps 0;  
Matches 174; Conservative 0; Mismatches 0;  
QY 620 REKNSKREBEEOQEGGFASPRGTGKWKRSAGGSPSAGSPRTSKTKTVEPYSLTAQ 679  
Db 1 REKNSKREBEEOQEGGFASPRGTGKWKRSAGGSPSAGSPRTSKTKTVEPYSLTAQ 60  
QY 680 QSSLIREDKSNKLNWNEVLASLKDPRPASGSPQLFLSKVEETFCICCOELVFRPITTV 739  
Db 61 QSSLIREDKSNKLNWNEVLASLKDPRPASGSPQLFLSKVEETFCICCOELVFRPITTV 120  
QY 740 QHNVCKDCLDRSFRAQVFCPCRYDLGRSYAMQVNPQLQTVLNQLRPGYNGR 793  
Db 121 QHNVCKDCLDRSFRAQVFCPCRYDLGRSYAMQVNPQLQTVLNQLRPGYNGR 174  
RESULT 10  
AAU16348  
ID AAU16348 standard; protein; 133 AA.  
XX  
AC AAU16348;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human novel secreted protein, Seq ID 1301.  
XX  
KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
KW preservative; antiproliferative.  
XX  
OS Homo sapiens.  
XX  
FN WO200155322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001341.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0246177P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-488783/53.  
N-PSDB; AAS26335.  
New nucleic acid molecules encoding 461 human secreted proteins for  
diagnosing, preventing, treating or ameliorating medical conditions and  
used as food additives or preservatives.  
Claim 11; SEQ ID NO 1301; 980pp; English.  
The invention relates to isolated nucleic acid molecules and their  
encoded secreted proteins. The nucleic acids and proteins are used to  
prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
in diagnosing a pathological condition or susceptibility to a  
pathological condition. Antibodies to the proteins can also be used in  
alleviating symptoms associated with the disorders and in diagnostic  
immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays  
(ELISA). Disorders which are diagnosed or treated include autoimmune  
diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
nervous system disorders e.g. Alzheimer's disease, infections caused by  
bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
and many other disorders listed in the specification. The polypeptides  
can also be used to aid wound healing and epithelial cell proliferation,  
to prevent skin aging due to sunburn, to maintain organs before  
transplantation, for supporting cell culture of primary tissues, to  
regenerate tissues and in chemotaxis. The polypeptides can also be used  
as a food additive or preservative to increase or decrease storage  
capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
minerals, cofactors and other nutritional components. The present  
sequence represents a novel secreted protein of the invention. Note: The  
sequence data for this patent did not form part of the printed

```

Query Match          10.8%; Score 86; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.6e-79;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 667 KKTKEPYSLTAAQSSLIREDKSNKLNVEVLASLKDRPASGPFQLFLSKVEETFCQIC 726
Db 7 KKTKEPYSLTAAQSSLIREDKSNKLNVEVLASLKDRPASGPFQLFLSKVEETFCQIC 66

QY 727 COELVFRPITTVCOHNVCKDCLDRSF 752
Db 67 COELVFRPITTVCOHNVCKDCLDRSF 92

RESULT 11
ABUS5417
ID ABUS5417 standard; protein; 133 AA.
XX
AC ABUS5417;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polypeptide #504.
XX
KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PN US2002132753-A1.
XX
PD 19-SEP-2002.
XX
PF 17-JAN-2001; 2001US-00764864.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 08-SEP-2000; 2000US-0229513P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.

```

```

PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-147444/14.
DR N-PSDB; ABX73676.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 11; SEQ ID NO 1301; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABUS4914-ABUS5699 and
CC ABUS5748 represent human novel polypeptides of the invention
XX
SQ Sequence 133 AA;

Query Match          10.8%; Score 86; DB 6; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.6e-79;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 667 KKTKEPYSLTAAQSSLIREDKSNKLNVEVLASLKDRPASGPFQLFLSKVEETFCQIC 726
Db 7 KKTKEPYSLTAAQSSLIREDKSNKLNVEVLASLKDRPASGPFQLFLSKVEETFCQIC 66

QY 727 COELVFRPITTVCOHNVCKDCLDRSF 752
Db 67 COELVFRPITTVCOHNVCKDCLDRSF 92

RESULT 12
ADM20128
ID ADM20128 standard; protein; 148 AA.
XX
AC ADM20128;
XX
DT 20-MAY-2004 (first entry)
XX

```

DE Protein encoded by novel human channel/transporter gene #206 clone 2.  
XX immunosuppressive; antiarthritic; antirheumatic; antiproliferative;  
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;  
KW neuroprotective; antibacterial; virucide; fungicide; opthalmological;  
KW gene therapy; channel/transporter protein; rheumatoid arthritis;  
KW neoplasm; cardiac arrest; cerebrovascular disorder; cerebral ischemia;  
KW angiogenesis; nervous system disorder; Alzheimer's disease;  
KW ocular disorder; corneal infection; wound healing;  
KW epithelial cell proliferation; skin aging; sunburn; transplantation;  
KW chemotaxis; food additive.  
XX Homo sapiens.  
OS  
XX  
PN WO200154472-A2.  
XX  
PN  
XX  
PD  
XX  
PF 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001307.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-022547P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229309P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-023935P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244517P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249255P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250316P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 DR WPI: 2001-476159/51.  
 DR N-PSDB; ADM19649.  
 XX  
 PT Isolated nucleic acid molecule encoding a channel/transporter protein is  
 PT used in preventing, treating or ameliorating a medical condition.  
 XX  
 PS Claim 11; SEQ ID NO 935; 809pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC channel/transporter protein or sequences at least 95% identical to a  
 CC these. The nucleic acids and proteins encoded by them are used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. The antibodies to the proteins can also be used  
 CC in alleviating symptoms associated with the disorders and in diagnostic  
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
 CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,  
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.  
 CC The polypeptides can also be used to aid wound healing and epithelial  
 CC cell proliferation, to prevent skin aging due to sunburn, to maintain  
 CC organs before transplantation, for supporting cell culture of primary  
 CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can  
 CC also be used as a food additive or preservative to increase or decrease  
 CC storage capabilities. This sequence corresponds to a protein of the  
 CC invention.  
 XX  
 SQ Sequence 148 AA;

Query Match 10.1%; Score 80; DB 4; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-73;  
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWIQVRTMDGROTHVDSLSRLTKVEELRRKIQELFHVPGQLRFLVRGKQMGDHTLFD 60  
 Dd |||||  
 QY 61 YEVRNDTIQLLVRSVLVP 80  
 Dd |||||  
 Dd 108 YEVRNDTIQLLVRSVLVP 127

RESULT 13  
 ADM19888  
 ID ADM19888 standard; protein; 150 AA.  
 XX  
 XX ADM19888;  
 AC  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Protein encoded by novel human channel/transporter gene #206.  
 DE immunosuppressive; antiarthritic; antirheumatic; antiproliferative;  
 XX cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;  
 KW

KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
 KW gene therapy; channel/transporter protein; rheumatoid arthritis;  
 KW neoplasm; cardiac arrest; cerebrovascular disorder; cerebral ischemia;  
 KW angiogenesis; nervous system disorder; Alzheimer's disease;  
 KW ocular disorder; corneal infection; wound healing;  
 KW epithelial cell proliferation; skin aging; sunburn; transplantation;  
 KW chemotaxis; food additive.

XX Homo sapiens.  
 OS  
 XX WO200154472-A2.  
 PN  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US0001307.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 05-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 12-SEP-2000; 2000US-0232081P.  
 PR 14-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.

```
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246509P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2000US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-476159/51.
XX N-PSDB; ADM19409.
XX Isolated nucleic acid molecule encoding a channel/transporter protein is
PT used in preventing, treating or ameliorating a medical condition.
XX Claim 11; SEQ ID NO 695; 809pp; English.
XX The invention relates to an isolated nucleic acid molecule encoding a
CC channel/transporter protein or sequences at least 95% identical to a
CC these. The nucleic acids and proteins encoded by them are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. The antibodies to the proteins can also be used
CC in alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. This sequence corresponds to a protein of the
CC invention.
XX SQ Sequence 150 AA;
Query Match 10.1%; Score 80; DB 4; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.7e-73;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWIQVITMDGROTHVTVDLSRLTKVEELRRKIQELPHVEPGLQRLFYRGKQMEDGHTLFD 60
DB 4B MWIQVITMDGROTHVTVDLSRLTKVEELRRKIQELPHVEPGLQRLFYRGKQMEDGHTLFD 107
QY 61 YEVRLNDTIQLLVROSLVLP 80
DB 108 YEVRLNDTIQLLVROSLVLP 127
RESULT 14
ABB76982
ID ABB76982 standard; peptide; 26 AA.
XX ABB76982;
AC ABB76982;
XX 22-JUL-2002 (first entry)
XX Human Inverted CCAAT box binding protein, ICBP90, fragment #2.
DE Human; inverted CCAAT box binding protein; ICBP90; cytosstatic;
XX cell proliferation control; inverted CCAAT box; cancer.
OS Homo sapiens.
XX WO200078949-A1.
PN
```

```

XX 28-DEC-2000.
XX
XX
XX 22-JUN-2000; 2000WO-FR001747.
XX
XX 22-JUN-1999; 99FR-00007935.
XX
XX (ADER-) ADEREGEM ASSOC DEV RECH EN GENETIQ.
XX
XX Bronner C, Hopfner R, Mousli M, Jeltsch J, Lutz Y, Oudet P;
XX
XX WPI; 2001-091571/10.
XX
XX N-PSDB; ABL58022.
XX
XX Novel inverted CCAAT box binding protein, and related nucleic acids,
XX antibodies and specific ligands, useful for treating and preventing
XX cancer.
XX
XX Claim 2; Page 102; 115pp; French.
XX
XX The present sequence is a peptide fragment of human ICBP90 (inverted
XX CCAAT box binding protein). The inverted CCAAT box is implicated in cell
XX proliferation control. Several copies of the inverted CCAAT box are
XX present in the promoter of the topoisomerase IIalpha gene, and also
XX functions as a nuclear receptor. ICBP90 (ABB76980) and its coding
XX sequence (ABL58020) are useful for treatment and/or prevention of cancer
XX
XX Sequence 26 AA;
SQ
Query Match 3.2%; Score 25; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 2.3e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 289 MVDNPMRRKSGPSCKHCKDQVNRLC 313
Db 1 MVDNPMRRKSGPSCKHCKDQVNRLC 25
|||||
RESULT 15
ABP64013
ID ABP64013 standard; protein; 198 AA.
AC
AC ABP64013;
XX
XX 04-NOV-2002 (first entry)
DT
DT
DE
DE Human ORF383.
XX
XX Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
XX Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
XX human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
XX cancer; cardiovascular disease; allergy; autoimmune disease;
XX wound healing; blood coagulation disorder; inflammatory disorder.
XX
XX Homo sapiens.
OS
XX
XX
XX US2002082206-A1.
XX
XX 27-JUN-2002.
XX
XX 30-MAY-2001; 2001US-00867550.
XX
XX 30-MAY-2000; 2000US-0208427P.
XX
XX (LEAC/) LEACH M D.
XX (MEHR/) MEHRABAN F.
XX (CONL/) CONLEY P B.
XX (TOPP/) TOPPER J N.
XX (LAND/) LAW D.
XX
XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
XX WPI; 2002-626554/67.
XX
DR N-PSDB; ABQ98576.
XX
XX New polypeptide designated ORFX are present in human atherogenic cells
XX and are useful to prevent and treat ORFX-associated disorders including
XX cancer, allergy, wound healing or autoimmune, cardiovascular or
XX inflammatory disease.
XX
XX Claim 10; SEQ ID NO 766; 78pp; English.
XX
XX The present invention relates to novel human ORFX polypeptides and their
XX coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
XX were discovered in human atherogenic cells, in particular in platelets
XX and human umbilical vein endothelial cells (HUVEC) and are expressed in
XX many other tissues as well. Atherogenic cells are cells which have the
XX potential to develop atherosclerotic plaques. The ORFX polypeptides and
XX nucleic acids are useful for treating or preventing a pathological
XX condition associated with an ORFX-associated disorder, e.g. cancer,
XX cardiovascular disease, allergy, autoimmune disease, wound healing, blood
XX coagulation disorders or inflammatory disorders. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from the USPTO web site at
XX seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
XX Sequence 198 AA;
SQ
Query Match 2.4%; Score 19; DB 5; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.4e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 409 ECTIVPSNHYGPIPGIPVG 427
Db 26 ECTIVPSNHYGPIPGIPVG 44
|||||
Search completed: November 1, 2004, 15:53:18
Job time : 95 secs

```